Moran, M.
09/779240

FILE 'HCAPLUS' ENTERED AT 11:12:11 ON 16 APR 2003
L1 1 S PREDICTOR SET AND COMPLEMENT?

L1 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:778756 HCAPLUS
DOCUMENT NUMBER: 137:274049
TITLE: Fast microarray expression data analysis method for network exploration and identifying the minimum number of markers defining a given state
INVENTOR(S): Kanevsky, Valery; Vailaya, Aditya
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 17 pp.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002147546	A1	20021010	US 2001-779240	20010208

PRIORITY APPLN. INFO.: US 2001-779240 20010208

AB A method for performing network reconstruction form microarray hybridization data is described. The method is used to identify genes showing coordinated regulation or membership of the same regulatory network. The method includes the steps of selecting a **predictor set** of features, adding a **complement** to the **predictor set** based on a quality of prediction, checking to see if all of the features of the **predictor set** are repeated, and then removing one feature from the **predictor set**. The algorithm and method repeats the steps of adding a **complement**, checking the **predictor set** and removing a feature until the features of the **predictor set** are repeated. If the features of the **predictor set** are repeated, the algorithm and method terminate.

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO, COMPENDEX, PROMT, INSPEC, COMPUSCIENCE' ENTERED AT 11:12:53 ON 16 APR 2003)
L2 5 S L1
L3 4 DUP REM L2 (1 DUPLICATE REMOVED)

L3 ANSWER 1 OF 4 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2003-237981 [23] WPIDS
DOC. NO. NON-CPI: N2003-189565
DOC. NO. CPI: C2003-060896
TITLE: Determination of **predictor set** of features associated with targets useful in bio-informatics, involves adding **complement** to selected **predictor set** based on quality of prediction.
DERWENT CLASS: B04 D16 S03 T01 T02
INVENTOR(S): KANEVSKY, V; VAILAYA, A
PATENT ASSIGNEE(S): (KANE-I) KANEVSKY V; (VAIL-I) VAILAYA A
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2002147546	A1	20021010	(200323)*		17

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2002147546	A1	US 2001-779240	20010208

PRIORITY APPLN. INFO: US 2001-779240 20010208
AN 2003-237981 [23] WPIDS
AB US2002147546 A UPAB: 20030407
NOVELTY - Determination (M1) of **predictor set** of features associated with targets comprising selecting a **predictor set** of features, adding a **complement**(s) in the **predictor set** based on a quality of prediction, and checking to see if all of the features are repeated, and removing a feature from the **predictor set**, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

(1) network reconstruction comprising selecting a target, selecting a **predictor set** of features, adding a **complement** to the **predictor set** based on a quality of prediction, checking to see if all of the features are repeated, and removing one feature from the **predictor set**;

(2) classification of experiments, comprising selecting a target whose values are chosen to represent the class of the experiment, and performing the network reconstruction;

(3) a computer software for performing network reconstruction, comprising an algorithm; and

(4) a system for performing network reconstruction comprising a computer and a computer software for running on the computer, the computer software performing network reconstruction.

USE - (M1) is useful for determining **predictor set** of features associated with a target useful in the field of bio-informatics. It can also be used in information, computer, software and data processing systems.

ADVANTAGE - (M1) is capable of identifying a good subset of genes that can be predicted the target vector. Hence, these genes can form a diagnostic set of genes whose expression values are used to discriminate between the various types of leukemia.

DESCRIPTION OF DRAWING(S) - The figure shows a schematic view of the invention in vector format illustrating the target and the predictors.

Dwg.1/8

L3 ANSWER 2 OF 4 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:298035 BIOSIS
DOCUMENT NUMBER: PREV200200298035
TITLE: Gene expression profiles of poor-prognosis primary breast cancer correlate with survival.
AUTHOR(S): Bertucci, Francois; Nasser, Valery; Granjeaud, Samuel; Eisinger, Francois; Adelaide, Jose; Tagett, Rebecca; Loriod, Beatrice; Giaconia, Aurelia;

Benziane, Athmane; Devilard, Elisabeth; Jacquemier, Jocelyne; Viens, Patrice; Nguyen, Catherine; Birnbaum, Daniel (1); Houlgatte, Remi

CORPORATE SOURCE: (1) Laboratoire d'Oncologie Moleculaire, U.119 Inserm, 27 Boulevard Lei Roure, IFR57, 13009, Marseille: birnbaum@marseille.inserm.fr France

SOURCE: Human Molecular Genetics, (15 April, 2002) Vol. 11, No. 8, pp. 863-872. http://hmg.oupjournals.org/. print.
ISSN: 0964-6906.

DOCUMENT TYPE: Article
LANGUAGE: English

AB The extensive heterogeneity of breast cancer complicates the precise assessment of tumour aggressiveness, making therapeutic decisions difficult and treatments inappropriate in some cases. Consequently, the long-term metastasis-free survival rate of patients receiving adjuvant chemotherapy is only 60%. There is a genuine need to identify parameters that might accurately predict the effectiveness of this treatment for each patient. Using cDNA arrays, we profiled tumour samples from 55 women with poor-prognosis breast cancer treated with adjuvant anthracycline-based chemotherapy. Gene expression monitoring was applied to a set of about 1000 candidate cancer genes. Differences in expression profiles provided molecular evidence of the clinical heterogeneity of disease. First, we confirmed the capacity of a 23-gene **predictor set** , identified in a previous study, to distinguish between tumours associated with different survival. Second, using a refined gene set derived from the previous one, we distinguished, among the 55 clinically homogeneous tumours, three classes with significantly different clinical outcome; 5-year overall survival and metastasis-free survival rates were respectively 100% and 75% in the first class, 65% and 56% in the second and 40% and 20% in the third. This discrimination resulted from the differential expression of two clusters of genes encoding proteins with diverse functions, including the estrogen receptor (ER). Another finding was the identification of two ER-positive tumour subgroups with different survival. These results indicate that gene expression profiling can predict clinical outcome and lead to a more precise classification of breast tumours. Furthermore, the characterization of discriminator genes might accelerate the development of new specific and alternative therapies, allowing more rationally tailored treatments that are potentially more efficient and less toxic.

L3 ANSWER 3 OF 4 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 1999:79819 PROMT
TITLE: Content Analysis as a Predictive Methodology: Recall, Readership, and Evaluations of Business-to-Business Print Advertising.
AUTHOR(S): A., John L.; Neuendorf Naccarato Kimberly
SOURCE: Journal of Advertising Research, (1 May 1998) .
ISSN: 0021-8499.
PUBLISHER: Advertising Research Foundation, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 6048

AB *FULL TEXT IS AVAILABLE IN THE ALL FORMAT*

JOHN L. NACCARATO

L3 ANSWER 4 OF 4 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
DUPLICATE 1

ACCESSION NUMBER: 1997:485098 BIOSIS
DOCUMENT NUMBER: PREV199799784301
TITLE: When are leaves good thermometers? A new case for Leaf Margin Analysis.
AUTHOR(S): Wilf, Peter
CORPORATE SOURCE: Dep. Paleobiol., MRC 121, Natl. Museum Natural History, Smithsonian Inst., Washington, DC 20560 USA
SOURCE: Paleobiology, (1997) Vol. 23, No. 3, pp. 373-390. ISSN: 0094-8373.
DOCUMENT TYPE: Article
LANGUAGE: English
AB Precise estimates of past temperatures are critical for understanding the evolution of organisms and the physical biosphere, and data from continental areas are an indispensable **complement** to the marine record of stable isotopes. Climate is considered to be a primary selective force on leaf morphology, and two widely used methods exist for estimating past mean annual temperatures from assemblages of fossil leaves. The first approach, Leaf Margin Analysis, is univariate, based on the positive correlation in modern forests between mean annual temperature and the proportion of species in a flora with untoothed leaf margins. The second approach, known as the Climate-Leaf Analysis Multivariate Program, is based on a modem data set that is multivariate. I argue here that the simpler, univariate approach will give paleotemperature estimates at least as precise as the multivariate method because (1) the temperature signal in the multivariate data set is dominated by the leaf-margin character; (2) the additional characters add minimal statistical precision and in practical use do not appear to improve the quality of the estimate; (3) the predictor samples in the univariate data set contain at least twice as many species as those in the multivariate data set; and (4) the presence of numerous sites in the multivariate data set that are both dry and extremely cold depresses temperature estimates for moist and nonfrigid paleofloras by about 2 degree C, unless the dry and cold sites are excluded from the **predictor set**. New data from Western Hemisphere forests are used to test the univariate and multivariate methods and to compare observed vs. predicted error distributions for temperature estimates as a function of species richness. Leaf Margin Analysis provides excellent estimates of mean annual temperature for nine floral samples. Estimated temperatures given by 16 floral subsamples are very close both to actual temperatures and to the estimates from the samples. Temperature estimates based on the multivariate data set for four of the subsamples were generally less accurate than the estimates from Leaf Margin Analysis. Leaf-margin data from 45 transect collections demonstrate that sampling of low-diversity floras at extremely local scales can result in biased leaf-margin percentages because species abundance patterns are uneven. For climate analysis, both modern and fossil floras should be sampled over an area sufficient to minimize this bias and to maximize recovered species richness within a given climate.

(FILE 'HCAPLUS' ENTERED AT 11:13:47 ON 16 APR 2003)

Searcher : Shears 308-4994

```
L4        10557 SEA FILE=HCAPLUS ABB=ON  PLU=ON  PREDICT? AND COMPLEMENT?

L5          743 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L4 AND (REPEAT? OR
                REPETIT?)
L6          722 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L5 AND (PROTEIN OR
                POLYPROTEIN OR PEPTIDE OR POLYPEPTIDE OR GENE OR DNA OR
                DEOXYRIBONUCLEIC OR DEOXY(W)RIBONUCLEIC OR NUCLEIC OR
                IMMUNOL?(W)(DATA OR INFORM?))
L7          157 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L6 AND (FEATURE OR
                CHARACTERIST? OR TRAIT OR QUALIT? OR MARKER)
L8           14 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L7 AND (ALGORITHM OR
                GSSA OR METHOD OR TECHNIQUE)

=> s l8 not l1
L9           13 L8 NOT L1

L9   ANSWER 1 OF 13  HCAPLUS  COPYRIGHT 2003 ACS
ACCESSION NUMBER:        2003:202899  HCAPLUS
DOCUMENT NUMBER:         138:232943
TITLE:                   Gene expression profiling in the
                         diagnosis and prognosis of cancer
INVENTOR(S):             Gordon, Gavin J.; Jensen, Roderick V.; Gullans,
                         Steven R.; Bueno, Raphael
PATENT ASSIGNEE(S):      The Brigham and Women's Hospital, Inc., USA
SOURCE:                  PCT Int. Appl., 396 pp.
                         CODEN: PIXXD2
DOCUMENT TYPE:           Patent
LANGUAGE:                English
FAMILY ACC. NUM. COUNT:  1
PATENT INFORMATION:

     PATENT NO.         KIND  DATE          APPLICATION NO.  DATE
     ---------------    ----  --------      ---------------  --------
     WO 2003021229       A2   20030313      WO 2002-US28203  20020905
         W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH,
             CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD,
             GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,
             LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
             NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ,
             TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW,
             AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
         RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE,
             BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU,
             MC, NL, PT, SE, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
             GW, ML, MR, NE, SN, TD, TG
PRIORITY APPLN. INFO.:                  US 2001-317389P  P  20010905
                                        US 2002-236031   A  20020830
```

AB Genes that are differentially expressed in normal and diseased tissue are identified for use in the diagnosis and prognosis of disease such as cancer, esp. pleural mesothelioma. Sets of genes that are expressed differentially in malignant pleural mesothelioma are reported. These sets of genes can be used to discriminate between normal and malignant tissues, and between classes of malignant tissues. Accordingly, diagnostic assays for classification of tumors, prediction of tumor outcome, selecting and monitoring treatment regimens and monitoring tumor progression/regression also are provided.

L9 ANSWER 2 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2003:37277 HCAPLUS
DOCUMENT NUMBER: 138:233653
TITLE: Structural modeling of ataxin-3 reveals distant homology to adaptins
AUTHOR(S): Albrecht, Mario; Hoffmann, Daniel; Evert, Bernd O.; Schmitt, Ina; Wullner, Ullrich; Lengauer, Thomas
CORPORATE SOURCE: Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Sankt Augustin, 53754, Germany
SOURCE: Proteins: Structure, Function, and Genetics (2003), 50(2), 355-370
CODEN: PSFGEY; ISSN: 0887-3585
PUBLISHER: Wiley-Liss, Inc.
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Spinocerebellar ataxia type 3 (SCA3) is a polyglutamine disorder caused by a CAG **repeat** expansion in the coding region of a **gene** encoding ataxin-3, a **protein** of yet unknown function. Based on a comprehensive computational anal., we propose a structural model and structure-based functions for ataxin-3. Our **predictive** strategy comprises the compilation of multiple sequence and structure alignments of carefully selected **proteins** related to ataxin-3. These alignments are consistent with addnl. information on sequence motifs, secondary structure, and domain architectures. The application of **complementary methods** revealed the homol. of ataxin-3 to ENTH and VHS domain **proteins** involved in membrane trafficking and regulatory adaptor functions. We modeled the structure of ataxin-3 using the adaptin AP180 as a template and assessed the reliability of the model by comparison with known sequence and structural **features**. We could further infer potential functions of ataxin-3 in agreement with known exptl. data. Our database searches also identified an as yet uncharacterized family of **proteins**, which we named josephins because of their pronounced homol. to the Josephin domain of ataxin-3.
REFERENCE COUNT: 185 THERE ARE 185 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 3 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:517470 HCAPLUS
DOCUMENT NUMBER: 138:50370
TITLE: Determining the genomic locations of **repetitive DNA** sequences with a whole-genome microarray: IS6110 in Mycobacterium tuberculosis
AUTHOR(S): Kivi, Marten; Liu, Xuemin; Raychaudhuri, Soumya; Altman, Russ B.; Small, Peter M.
CORPORATE SOURCE: Division of Infectious Diseases and Geographic Medicine, Stanford University, Stanford, CA, 94305, USA
SOURCE: Journal of Clinical Microbiology (2002), 40(6), 2192-2198
CODEN: JCMIDW; ISSN: 0095-1137
PUBLISHER: American Society for Microbiology

DOCUMENT TYPE: Journal
LANGUAGE: English
AB The mycobacterial insertion sequence IS6110 has been exploited extensively as a clonal **marker** in mol. epidemiol. studies of tuberculosis. In addn., it has been hypothesized that this element is an important driving force behind genotypic variability that may have phenotypic consequences. The authors present here a novel, **DNA** microarray-based methodol., designated SiteMapping, that simultaneously maps the locations and orientations of multiple copies of IS6110 within the genome. To investigate the sensitivity, accuracy, and limitations of the **technique**, it was applied to eight Mycobacterium tuberculosis strains for which complete or partial IS6110 insertion site information had been detd. previously. SiteMapping correctly located 64% (38 of 59) of the IS6110 copies **predicted** by restriction fragment length polymorphism anal. The **technique** is highly specific; 97% of the **predicted** insertion sites were true insertions. Eight previously unknown insertions were identified and confirmed by PCR or sequencing. The performance could be improved by modifications in the exptl. protocol and in the approach to data anal. SiteMapping has general applicability and demonstrates an expansion in the applications of microarrays that **complements** conventional approaches in the study of genome architecture.
REFERENCE COUNT: 32 THERE ARE 32 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 4 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:57053 HCAPLUS
DOCUMENT NUMBER: 132:291914
TITLE: The flow cytometric pattern of CD34, CD15 and CD13 expression in acute myeloblastic leukemia is highly **characteristic** of the presence of PML-RAR.alpha. **gene** rearrangements
AUTHOR(S): Orfao, Alberto; Chillon, Maria Carmen; Bortoluci, Aglae M.; Lopez-Berges, Maria Consuelo; Garcia-Sanz, Ramon; Gonzalez, Marcos; Tabernero, Maria Dolores; Garcia-Marcos, Maria Antonia; Rasillo, Ana Isabel; Hernandez-Rivas, Jesus; Miguel, Jesus F. San
CORPORATE SOURCE: Servicio General de Citometna, Centro de Investigaciones del Cancer de Salamanca and Departamento de Medicina, Universidad de Salamanca, Salamanca, Spain
SOURCE: Haematologica (1999), 84(5), 405-412
CODEN: HAEMAX; ISSN: 0390-6078
PUBLISHER: Il Pensiero Scientifico Editore
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Rapid identification of AML patients carrying the t(15;17) translocation for treatment decision-making is currently made on the basis of morphol. screening. However, the existence of both false positives and negatives highlights the need for more objective **methods** of screening AML cases and further mol. confirmation of the t(15;17) translocation. Here, a total of 111 AML cases were analyzed to investigate whether immunophenotyping based on the

assessment of multiple-stainings analyzed at flow cytometry could improve the sensitivity and specificity of morphol. identification of acute promyelocytic leukemia (APL) carrying the t(15;17) translocation. FISH anal. was used as a **complementary technique** for cases in which morphol. and mol. biol. yielded discrepant results. Concordant results between morphol. and RT-PCR were found in 102/111 (91.8%) cases: 34 patients had M3/PML-RAR.alpha.+ and 68 non-M3/PML-RAR.alpha.- disease. Nine cases showed discrepants results. Multivariate anal. showed that the best combination of immunol. **markers** for discriminating between M3/PML-RAR.alpha.+ and non-M3/PML-RAR.alpha.- cases was that of the presence of heterogeneous expression of CD13, the existence of a single major blast-cell population, and a **characteristic** CD34/CD15 phenotypic pattern. A score system based on these parameters was designed, and the 34 M3/PML-RAR.alpha.+ cases showed a score of 3 (presence of the 3 phenotypic **characteristics**). In contrast, only 1 out of the 68 (1.3%) non-M3/PML-RAR.alpha.- cases had this score, most of these latter cases (53/68, 78%) scoring either 0 or 1. Therefore, among these cases, immunophenotyping showed a sensitivity of 100% and a specificity of 99% for **predicting** PML/RAR.alpha. **gene** rearrangements. Of the 9 cases in which morphol. and mol. biol. results were discrepant, four cases displayed M3 morphol. without PML/RAR.alpha. rearrangements by RT-PCR. In only one of these 4 cases did the immunophenotype score 3, this being the only FISH-pos. case. From the remaining five discrepant cases (non-M3 morphol. while pos. for PML/RAR.alpha.), two cases had a phenotypic score of 3 and were FISH-pos., whereas the other three were neg. by FISH. Upon **repeating** RT-PCR studies, two of these latter three cases became neg. Thus, immunophenotyping may be of great value for quick screening of APL with PML/RAR.alpha. rearrangements.

REFERENCE COUNT: 42 THERE ARE 42 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 5 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:256014 HCAPLUS
DOCUMENT NUMBER: 124:336629
TITLE: Molecular cloning and characterization of a novel human diacylglycerol kinase .zeta.
AUTHOR(S): Bunting, Michaeline; Tang, Wen; Zimmerman, Guy A.; McIntyre, Thomas M.; Prescott, Stephen M.
CORPORATE SOURCE: Eccles Program Human Mol. Biol. Genet., Univ. Utah, Salt Lake City, UT, 84112, USA
SOURCE: Journal of Biological Chemistry (1996), 271(17), 10230-6
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Diacylglycerol (DAG) occupies a central position in the synthesis of complex lipids and also has important signaling roles. For example, DAG is an allosteric regulator of **protein** kinase C, and the cellular levels of DAG may influence a variety of processes including growth and differentiation. The authors previously demonstrated that human endothelial cells derived from umbilical vein express growth-dependent changes in their basal levels of

diacylglycerol and diacylglycerol kinase activity (Whatley, R. E., Stroud, E. D., Bunting, M., Zimmerman, G. A., McIntyre, T. M., and Prescott, S. M. (1993) J. Biol. Chem. 268, 16130-16138). To further explore the role of diacylglycerol metab. in endothelial responses, the authors used a degenerate reverse transcription-polymerase chain reaction **method** to identify diacylglycerol kinase isoenzymes expressed by human endothelial cells. They report the isolation of a 3.5-kilobase cDNA encoding a novel diacylglycerol kinase (hDGK.zeta.) with a **predicted** mol. mass of 103.9 kDa. Human DGK.zeta. contains two zinc fingers, an ATP binding site, and four ankyrin **repeats** near the carboxyl terminus. A unique **feature**, as compared with other diacylglycerol kinases, is the presence of a sequence homologous to the MARCKS phosphorylation site domain. From Northern blot anal. of multiple tissues, it was obsd. that hDGK.zeta. mRNA is expressed at highest levels in brain. COS-7 cells transfected with the hDGK.zeta. and cDNA express 117-kDa and 114-kDa **proteins** that react specifically with an antibody to a **peptide** derived from a unique sequence in hDGK.zeta.. The transfected cells also express increased diacylglycerol kinase activity, which is not altered in the presence of R59949, an inhibitor of human platelet DGK activity. The hDGK.zeta. displays stereoselectivity for 1,2-diacylglycerol species in comparison to 1,3-diacylglycerol, but does not exhibit any specificity for mol. species of long chain diacylglycerols.

L9 ANSWER 6 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:21537 HCAPLUS
DOCUMENT NUMBER: 124:45913
TITLE: Structural **predictions** for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions
AUTHOR(S): Jiang, Xuliang; Dreano, Michel; Buckler, David R.; Cheng, Shirley; Ythier, Arnaud; Wu, Hao; Hendrickson, Wayne A.; Tayar, Nabil El
CORPORATE SOURCE: Howard Hughes Med. Inst., Columbia Univ., New York, NY, 10032, USA
SOURCE: Structure (London) (1995), 3(12), 1341-53
PUBLISHER: CODEN: STRUE6; ISSN: 0969-2126
DOCUMENT TYPE: Current Biology
LANGUAGE: Journal
English

AB Glycoprotein hormones influence the development and function of the ovary, testis and thyroid by binding to specific high-affinity receptors. The extra-cellular domains of these receptors are members of the leucine-rich **repeat** (LRR) **protein** superfamily and are responsible for the high-affinity binding. The crystal structure of a glycoprotein hormone, namely human choriogonadotropin (hCG), is known, but neither the receptor structure, mode of hormone binding, nor mechanism for activation, have been established. Despite very low sequence similarity between exon-demarcated LRRs in the receptors and the LRRs of porcine RNase inhibitor (RI), the secondary structures for the two **repeat** sets are alike. Constraints on curvature and .beta.-barrel geometry from the sequence pattern for **repeated** .beta..alpha. units suggest that the receptors contain three-dimensional structures similar to that of RI. With the RI crystal structure as a template, models were constructed for exons 2-8 of the receptors. The model

for this portion of the choriogonadotropin receptor is **complementary** in shape and electrostatic **characteristics** to the surface of hCG at an identified focus of hormone-receptor interaction. The **predicted** models for the structures and mode of hormone binding of the glycoprotein hormone receptors are to a large extent consistent with currently available biochem. and mutational data. **Repeated** sequences in .beta.-barrel **proteins** are shown to have general implications for constraints on structure. Averaging **techniques** used here to recognize the structural motif in these receptors should also apply to other **proteins** with **repeated** sequences.

L9 ANSWER 7 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1995:265485 HCAPLUS
DOCUMENT NUMBER: 122:50623
TITLE: Identification of interactive sites of **proteins** and **protein** receptors by computer-assisted searches for **complementary peptide** sequences
AUTHOR(S): Fassina, Giorgio; Melli, Marialuisa
CORPORATE SOURCE: Protein Engineering, TECNOGEN S.C.p.A., Caserta, 81015, Italy
SOURCE: ImmunoMethods (1994), 5(2), 114-20
CODEN: IMUME8; ISSN: 1058-6687
DOCUMENT TYPE: Journal
LANGUAGE: English

AB **Protein** sites important for receptor binding have been identified in several systems by searching for **protein** /receptor stretches characterized by hydropathic anti-**complementarity**. A computer-assisted **method** [SITESEARCH] has been developed to identify **protein** sites responsible for receptor recognition, once the amino acid sequences of the **protein** ligand and its receptor are available. The software first dets. the hydropathic profiles of the two **polypeptide** chains under investigation, and then compares profiles of preselected length, detg. at the same time the degree of their hydropathic **complementarity**. The procedure is **repeated** until all the different segments in the two chains are compared. Fragments characterized by the maximal level of hydropathic **complementarity** are selected as putative binding sites. This approach has been initially applied to the interleukin-1.beta. (IL-1.beta.)/receptor type I case. SITESEARCH identified residues 88-99 in IL-1.beta. and 151-162 in the receptor as the sequence pair characterized by the maximal level of hydropathic **complementarity**. These fragments, once produced by chem. synthesis, have displayed specific recognition properties for each other, as detected by solid-phase binding assays. The IL-1.beta. sequence identified corresponds to a highly exposed part of the **protein** mol., and substitution of this sequence with another of the same length but with different hydropathic **characteristics** generated mutants with drastically reduced binding activity to the receptor. Mutations in this sequence did not alter the **protein** biol. activity, thus suggesting the structural integrity of the mutants. Cumulatively, these results validate the SITESEARCH **prediction**, suggesting that IL-1.beta. sequence 88-99 is

involved in at least a portion of the **protein**/receptor binding site.

L9 ANSWER 8 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:264014 HCAPLUS
DOCUMENT NUMBER: 120:264014
TITLE: Chicken filensin: A lens fiber cell **protein** that exhibits sequence similarity to intermediate filament **proteins**
AUTHOR(S): Remington, Susann G.
CORPORATE SOURCE: Dep. Genet. Cell Biol., Univ. Minnesota, St Paul, MN, 55108, USA
SOURCE: Journal of Cell Science (1993), 105(4), 1057-68
CODEN: JNCSAI; ISSN: 0021-9533
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Filensin, a 100 kDa, membrane-assocd., cytoskeletal **protein**, is uniquely expressed in the lens fiber cell (Merdes, A., et al., 1991, J. Cell Biol. 115: 397-410). The author cloned and sequenced a full-length chicken lens cDNA encoding filensin, also known as CP95 (Ireland, M. and Maisel, H. (1989) Lens and Eye Toxicity Research 6, 623-638). The deduced amino acid sequence of 657 residues contained an internal 280 residue heptad **repeat** domain with sequence similarities to the rod domain of intermediate filament **proteins**. The putative filensin rod domain could be divided into three .alpha.-helical segments (1A, 1B and 2) sepd. by short, non-helical linkers. The sequence of the amino-terminal end of the filensin rod domain contained the highly conserved intermediate filament segment 1A motif (Conway, J. F. and Parry, D. A. D. (1988) Int. J. Biol. Macromol. 10, 79-98). Allowing conservative amino acid substitutions, the sequence of the carboxy-terminal end of the filensin rod domain was similar to that of the highly conserved intermediate filament rod carboxy terminus. The .alpha.-helical segments of the shorter filensin rod domain aligned with the corresponding segments of intermediate filament **proteins** by allowing a gap of four heptad **repeats** in the amino-terminal half of filensin segment 2. Filensin rod segment 2 contained the **characteristic** stutter in heptad **repeat** phasing, nine heptads from the end of the intermediate filament rod. The overall sequence identity between the rod domains of filensin and individual intermediate filament **proteins** was 20 to 25%, approx. the level of sequence identity obsd. between intermediate filament **proteins** of different types. The open reading frame of chicken filensin **predicted** a 657 amino acid **protein** with mol. mass of 76 kDa. Embryonic chicken filensin migrated in SDS-PAGE as a triplet of 102, 105 and 109 kDa, while rooster filensin migrated as a 105 and 109 kDa doublet. Antibodies to filensin labeled lens fiber cells but not lens epithelial cells. By immunofluorescence **methods** filensin was localized to the fiber cell plasma membranes, including the ends of elongated fiber cells.

L9 ANSWER 9 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:3349 HCAPLUS
DOCUMENT NUMBER: 120:3349
TITLE: eps15, A novel tyrosine kinase substrate, exhibits transforming activity

AUTHOR(S): Fazioli, Francesca; Minichiello, Liliana; Matoskova, Brona; Wong, William T.; Di Fiore, Piere Paolo
CORPORATE SOURCE: Lab. Cell. Mol. Biol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
SOURCE: Molecular and Cellular Biology (1993), 13(9), 5814-28
CODEN: MCEBD4; ISSN: 0270-7306
DOCUMENT TYPE: Journal
LANGUAGE: English

AB An expression cloning **method** which allows direct isolation of cDNAs encoding substrates for tyrosine kinases was applied to the study of the epidermal growth factor (EGF) receptor (EGFR) signaling pathway. A previously undescribed cDNA was isolated and designated eps15. The structural **features** of the **predicted** eps15 **gene** product allow its subdivision into 3 domains. Domain I contains signatures of a regulatory domain, including a candidate tyrosine phosphorylation site and EF-hand-type calcium-binding domains. Domain II presents the **characteristic** heptad **repeats** of coiled-coil rod-like **proteins**, and domain III displays a **repeated** aspartic acid-proline-phenylalanine motif similar to a consensus sequence of several methylases. Antibodies specific for the eps15 **gene** product recognize 2 **proteins**: a major species of 142 kDa and a minor component of 155 kDa, both of which are phosphorylated on tyrosine following EGFR activation by EGF in vivo. EGFR is also able to directly phosphorylate the eps15 product in vitro. In addn., phosphorylation of the eps15 **gene** product in vivo is relatively receptor specific, since the erbB-2 kinase phosphorylates it very inefficiently. Finally, overexpression of eps15 is sufficient to transform NIH 3T3 cells, thus suggesting that the eps15 **gene** product is involved in the regulation of mitogenic signals.

L9 ANSWER 10 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1993:618813 HCAPLUS
DOCUMENT NUMBER: 119:218813
TITLE: Dinucleotide **repeat** polymorphism at the D11S982E locus
AUTHOR(S): Xiao, Hong; Ide, Susan E.; Merril, Carl R.; Polymeropoulos, Mihael H.
CORPORATE SOURCE: Neurosci. Cent., St. Elizabeths Hosp., Washington, DC, 20032, USA
SOURCE: Human Molecular Genetics (1993), 2(7), 1081
CODEN: HMGEE5; ISSN: 0964-6906
DOCUMENT TYPE: Journal
LANGUAGE: English

AB A (AC)n dinucleotide **repeat** sequence was isolated from a human fetal retinal cDNA library by hybridization with a synthetic poly (dC-dA).cntdot.(dG-dT) sequence. The insert of the polymorphic clone SE68 contg. the dinucleotide **repeat** was sequenced by the dideoxy **method**. The **predicted** length of the fragment amplified by PCR primers derived from the insert sequence was 121 bp. The frequencies of 5 alleles were estd. from 120 chromosomes of unrelated individuals. The dinucleotide **repeat marker** was assigned to chromosome 11 using 2 human-rodent somatic cell hybrid panels. Co-dominant segregation was obsd. in 5 informative families.

L9 ANSWER 11 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1993:228483 HCAPLUS
DOCUMENT NUMBER: 118:228483
TITLE: The iron-responsive element binding **protein**. Purification, cloning, and regulation in rat liver
AUTHOR(S): Yu, Yang; Radisky, Evette; Leibold, Elizabeth A.
CORPORATE SOURCE: Program Hum. Mol. Biol. Genet., Univ. Utah, Salt Lake City, UT, 84112, USA
SOURCE: Journal of Biological Chemistry (1992), 267(26), 19005-10
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The rat liver iron-responsive element binding **protein** (IRE-BP) was purified to homogeneity by chromatog. **methods** and partial amino acid sequence was obtained. A cDNA was isolated from a rat liver cDNA library and sequenced. The amino acid sequence deduced from the cDNA sequence corresponds to a **protein** of 889 amino acids with a **predicted** mol. wt. of 97,946. The NH2-terminal sequence obtained by Edman degrdn. matched the deduced amino acid sequence obtained from the cDNA, confirming the translational start site. Rat liver IRE-BP shares 95% identity with human IRE-BP and 98% identity with mouse IRE-BP indicating that the IRE-BPs have remained highly conserved during evolution. The 5'-untranslated region is at least 236 nucleotides and contains interesting structural **features** including two direct **repeats**, an inverted **repeat**, and three small open reading frames. The rat IRE-BP mRNA is approx. 3600 nucleotides and is expressed in a variety of rat tissues including liver, spleen, and gut. Over the course of 16 h following an i.p. injection of iron in rats, IRE-BP RNA binding activity decreases to 50% of control levels. The decrease in IRE-BP RNA binding activity in exts. from iron-treated rats is reversible by pretreatment of the exts. with reducing agents. The steady-state levels of IRE-BP mRNA remain const. during iron treatment. These data suggest that the decrease in IRE-BP RNA binding activity by iron in rat liver is due to post-translational changes in the RNA binding affinity of the IRE-BP and not due a decrease in the transcription of the IRE-BP **gene** or to the destabilization of the IRE-BP mRNA.

L9 ANSWER 12 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1993:227264 HCAPLUS
DOCUMENT NUMBER: 118:227264
TITLE: Dinucleotide **repeat** polymorphism at the D14S99E locus
AUTHOR(S): Polymeropoulos, Mihael H.; Xiao, Hong; Ide, Susan E.; Merril, Carl R.
CORPORATE SOURCE: Neurosci. Cent., Natl. Inst. Ment. Health, Washington, DC, 20032, USA
SOURCE: Human Molecular Genetics (1993), 2(4), 490
CODEN: HMGEE5; ISSN: 0964-6906
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The (AC)n dinucleotide **repeat** sequence was isolated from a human fetal retinal cDNA library by hybridization with a synthetic poly(dC-dA).cntdot.(dG-dT) sequence. The polymorphic clone was

designated SE60. The insert contg. the dinucleotide **repeat** was sequenced using the dideoxy **method.** The **predicted** length of the amplified fragment was 150 bp (GenBank accession no. L04461). 9 Alleles were detected in 142 chromosomes of unrelated individuals. The dinucleotide **repeat marker** was assigned to chromosome 14 using two human-rodent somatic cell hybrid panels. Co-dominant segregation was obsd. in five informative families.

L9 ANSWER 13 OF 13 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1982:63664 HCAPLUS
DOCUMENT NUMBER: 96:63664
TITLE: Enhanced graphic matrix analysis of **nucleic** acid and **protein** sequences
AUTHOR(S): Maizel, Jacob V., Jr.; Lenk, Robert P.
CORPORATE SOURCE: Lab. Mol. Genet., Natl. Inst. Child Health Hum. Dev., Bethesda, MD, 20205, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (1981), 78(12), 7665-9
CODEN: PNASA6; ISSN: 0027-8424
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The enhanced graphic matrix procedure analyzes **nucleic** acid and amino acid sequences for **features** of possible biol. interest and reveals the spatial patterns of such **features.** When a sequence is compared to itself, the **technique** shows regions of self-**complementarity,** direct **repeats,** and palindromic subsequences. Comparison of 2 different sequences, exemplified by Ig .kappa. light chain **genes,** by using colored graphic matrices showed domains of similarity, regions of divergence, and **features** explainable by transpositions. Anal. of mouse const. domain Ig sequences revealed self-**complementary** regions that can be used to fold the mol. into a structure consistent with electron microscopic observations. Computer translation of **nucleic** acid sequences into all possible amino acid sequences, followed by graphic matrix anal., provides a way to detect the most likely **protein** encoding regions and can **predict** the correct reading frames in sequences in which splicing patterns are not defined. Application of this **technique** to regions of simian virus 40 and polyoma virus demonstrates the frames of translation and shows the agreement of sequences detd. in sep. labs. with different virus isolates. The graphic matrix **technique** can also be used to assemble fragmentary sequences during detn., to display local variations in base compn., to detect distant evolutionary relations, and to display intragenic variation in rates of evolution.

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO, COMPENDEX, PROMT, INSPEC, COMPUSCIENCE' ENTERED AT 11:17:04 ON 16 APR 2003)
L10 172 S L8
L11 34 S L10 AND TARGET?
L12 33 S L11 NOT L2
L13 33 DUP REM L12 (0 DUPLICATES REMOVED)

L13 ANSWER 1 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2003:45823 PROMT
TITLE: Healthy hair, healthy segment. (Global Report: Hair Care).(Industry Overview)
SOURCE: Global Cosmetic Industry, (Jan 2003) Vol. 171, No. 1, pp. 37(14).
ISSN: ISSN: 1523-9470.
PUBLISHER: Allured Publishing Corp.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 5685
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Whether buying in a niche or mass market, store or salon, consumers want their hair to be an outward expression of themselves. And they are looking to you to help them scream style with healthy, shiny, sophisticated hair.
THIS IS THE FULL TEXT: COPYRIGHT 2003 Advanstar Communications, Inc.

Subscription: $40.00 per year. Published monthly. 362 S. Schamale Rd., Carol Stream, IL 60188-2787. FAX 708-653-2192.

L13 ANSWER 2 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2002:123625 PROMT
TITLE: Compugen Expands Proteomics Product Line; Introduces the Z4000 System for Management and Analysis of Large Scale 2-D Gel Experiments.
SOURCE: Business Wire, (26 Feb 2002) pp. 132.
PUBLISHER: Business Wire
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 706
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Business/High-Tech Editors, Health/Medical Writers
THIS IS THE FULL TEXT: COPYRIGHT 2002 Business Wire

L13 ANSWER 3 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2002:478658 PROMT
TITLE: Agency Profiles.
SOURCE: New Media Age, (19 Sep 2002) pp. S13(21).
ISSN: ISSN: 1364-7776.
PUBLISHER: Centaur Publishing Ltd.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 25777
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB 1. CONCHANGO
THIS IS THE FULL TEXT: COPYRIGHT 2002 Centaur Publishing Ltd.

Subscription: 149.00 British pounds per year. Published weekly. 50 Poland St., London, England W1V 4AX., United Kingdom

L13 ANSWER 4 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2002-500214 [53] WPIDS
DOC. NO. CPI: C2002-141658

TITLE: System for dispensing nanoliter sized droplets in defined distribution pattern to form miniarrays comprises print head with pipette-based dispensers, robotic arm for carrying print head and working platform.

DERWENT CLASS: B04 D16 P75 Q39

INVENTOR(S): SHAFER, D A

PATENT ASSIGNEE(S): (SHAF-I) SHAFER D A; (GENE-N) GENETAG TECHNOLOGY INC

COUNTRY COUNT: 87

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002040634	A2	20020523	(200253)*	EN	63
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW					
W: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZA ZW					
US 2002074342	A1	20020620	(200253)		
AU 2002035128	A	20020527	(200261)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002040634	A2	WO 2001-US43918	20011114
US 2002074342	A1 Provisional	US 2000-248247P	20001114
		US 2001-992516	20011114
AU 2002035128	A	AU 2002-35128	20011114

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2002035128	A Based on	WO 200240634

PRIORITY APPLN. INFO: US 2000-248247P 20001114; US 2001-992516 20011114

AN 2002-500214 [53] WPIDS

AB WO 200240634 A UPAB: 20020820

NOVELTY - System (S) for dispensing nanoliter-sized droplets on surface in precise pattern of non-overlapping spots to form two-dimensional miniarray assay comprising spotter device (I) with print head (Ia) and pipette-based dispensers (Ib), a robotic/mechanical arm carrying (Ia) and working platform for holding miniarray substrates or loading samples, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for:

(1) forming (M1) a miniarray where each known location or spot in the miniarray contains an analyte specific reagent for detecting an analyte in a sample involves:

(a) aspirating a solution of each analyte specific reagent with (Ib),

(b) pressuring a small defined droplet of analyte specific reagent from the narrow opening of the tip of (Ib),

(c) touching the droplet to the surface of the miniarray substrate with an action effective to release the droplet, thereby spotting a specific location in the miniarray with a specific volume of the analyte specific reagent, and

(d) **repeating** steps (a)-(c) until the miniarray is fabricated;

(2) diagnosing (M2, M3) a specific tissue or condition using specialized diagnostic miniarrays **targeted** to the analysis of the tissue or condition.

USE - (M2) or (M3) are useful for diagnosing a condition using specialized diagnostic miniarrays **targeted** to the analysis of the tissue or condition such as cancer, responses to an infection, responses to a therapeutic or toxic agent or stages of ageing (claimed). The miniarrays are useful in multi-analyte biological assays and are particularly useful for assessing **gene** expression profiles based on spotting cDNA or synthetic oligonucleotide samples.

ADVANTAGE - The miniarray formed by (M1) achieves a smaller, more condensed distribution by interspersing successive dispensing of reagents onto the array in regions between the spots dispensed previously. By shifting low density or larger format miniarrays, the development and use of low-resolution scanners is enabled. In addition such low density or large format miniarrays further enable the use of less sensitive labeling agents such as simple colored dyes as compared to the present need for expensive, non-permanent fluorescent labeling agents. The new miniarray instrument format enables ready customization of chips for the research and diagnostics market, eliminates problems of variation and expenses associated with miniaturized equipment, and facilitates development of small inexpensive instruments that can be more widely available for expression analysis. Thus, the miniarray can replace expensive microarrays with cheaper, larger format miniarrays with similar or equivalent diagnostic value. (M1) can also be used to create miniarrays on simple, small format substrate. The novel diagnostic miniarrays are specifically planned and spotted with disease or condition specific patterns built into their organization or arrangement. The presence of **gene** activity levels **predicted** for a specific disease, tissue or condition, such as up-regulated, down-regulated or unchanged activity levels, will create a simple recognizable clustered pattern in the array. This new diagnostic invention is enabled by the greater versatility and lower costs.

DESCRIPTION OF DRAWING(S) - The figure shows the mechanically operated pipetter, hydraulic pipetter and air driven pipetter operated remotely by microtubing. The hydraulic or air driven pipetters are typically actuated by stepper motor-driven syringe pumps.

Dwg.1A/9

L13 ANSWER 5 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2002-583343 [62] WPIDS
DOC. NO. CPI: C2002-164770
TITLE: Incorporating adenoviral **nucleic** acid into vector backbone, by homologous recombination of first adenoviral sequence and second, third adenoviral sequences of linear vector having non adenoviral sequence in host cell.
DERWENT CLASS: B04 D16

Searcher : Shears 308-4994

INVENTOR(S): YOUIL, R
PATENT ASSIGNEE(S): (MERI) MERCK & CO INC
COUNTRY COUNT: 22
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002031170	A1	20020418	(200262)*	EN	30

RW: AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
W: CA JP US

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002031170	A1	WO 2001-US42529	20011005

PRIORITY APPLN. INFO: US 2000-239547P 20001010
AN 2002-583343 [62] WPIDS
AB WO 200231170 A UPAB: 20020926
NOVELTY - Incorporating (M1) a **nucleic** acid sequence encoding adenovirus into vector backbone, comprising introducing a first adenoviral **nucleic** acid sequence (AS) (S1) encoding wild-type, mutant or replication-defective adenovirus, and linear vector having second (S2) and third (S3) ASs and a segment (S4) of non-AS, into a host cell, where homologous recombination of S1 with S2 and S3 results in vector with S4 and S1, is new.

DETAILED DESCRIPTION - Incorporating (M1) a **nucleic** acid sequence encoding adenovirus into vector backbone, comprising introducing S1 and a linear vector (I) having second (S2) and third (S3) adenoviral **nucleic** acid sequences and a segment of non-adenoviral **nucleic** acid sequence (S4) into a host cell, where S1 undergoes homologous recombination with S2 and S3 in the linear vector resulting in a vector comprising both the non-adenoviral segment of the linear vector and the first adenoviral sequence. S2 comprises basepairs corresponding to basepairs 1-100 of a wildtype adenovirus sequence not to extend past base pair 500 of a wildtype adenovirus sequence, and S3 comprises basepairs corresponding to basepairs 35835-35935 of a wildtype adenovirus sequence.

An INDEPENDENT CLAIM is also included for preparing, within a vector backbone, recombinant adenovirus carrying a desired **gene**, by:

(a) generating recombinant adenovirus carrying a desired **gene** in **complementation** cells by transfecting the cells with a vector comprising the desired **gene** flanked by adenoviral sequences homologous to a region within an adenoviral genome also present within the cells which is **targeted** for homologous recombination, extracting the recombinant adenoviral **nucleic** acid, and carrying out M1; or

(b) the **method** of:

(i) introducing S1 into **complementation** cells, where S1 comprises a restriction enzyme site within a region **targeted** for homologous recombination;

(ii) generating recombinant adenovirus carrying a desired **gene** in the **complementation** cells, by transfecting the cells with a vector comprising the desired **gene**

flanked by adenoviral sequences homologous to the region **targeted** for homologous recombination within the adenoviral genome;

(iii) digesting resultant cell products with a restriction enzyme capable of digesting the restriction enzyme site;

(iv) extracting the recombinant adenovirus from the **complementation** cells; and

(v) introducing into the bacterial cell both (I) and the recombinant adenovirus, where S1 undergoes homologous recombination with S2 and S3, resulting in a vector comprising S1 and S4.

USE - M1 is for incorporating a **nucleic** acid sequence encoding adenovirus into a vector backbone (claimed). M2 is useful for preparing a recombinant adenovirus carrying a desired **gene** within a vector backbone (claimed), preferably a circularized adenovirus. The **method** is also useful for isolating and analyzing variants present within a viral preparation, including those which are incapable of forming plaques. The **method** is also useful for producing plasmid preparation that are subsequently unsed in transient transfection studies to analyze transgene expression. The isolated cones may also be used in rescue experiments to identify whether or not these genomes are capable of forming infectious virus.

ADVANTAGE - The **method** produces a plasmid which is more readily analyzed by restriction digestion, polymerase chain reaction, **DNA** sequencing, or used in transient transfection studies. The adenovirus plasmids that are generated can be rescued back into virus form. The entire procedure took four days or less instead of the weeks required for plaque purification or dilution cloning isolating **techniques**. The **method** does not require the use of tissue culture materials or facilities. The **method** also allows for a more extensive and thorough examination of viral preparation, in that it allows for the detection of variants incapable of propagation without the assistance of coinfecting intact adenoviral genomes. Under standard conditions of plaque purification the variant genomes are not detected. It is **predicted** that far more variant genomes will be observed using the rapid **method** than would otherwise be detected by standard plaque purification **method** . The adenovirus is more readily manipulated and easier to screen. A large number of samples can be handled and processed easily. This amounts to a significant time and cost advantage. Large amounts of each adenoviral genome may be produced by cloning Escherichia coli, so that analysis is made by simple ethidium bromide staining after gel electrophoresis, rather than the more previous radiolabelling **techniques** used to visualize viral **DNA**. The adenoviral terminal portions utilized on the plasmid encompass much shorter regions of the adenoviral ends than previous vectors. This relieves the carry constrains of the importing plasmid, and the range of mutants capable of isolation through homologous recombination is broadened.

Dwg.0/4

L13 ANSWER 6 OF 33 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:516198 BIOSIS
DOCUMENT NUMBER: PREV200200516198
TITLE: Ahi-1, a novel **gene** encoding a modular **protein** with WD40-**repeat** and SH3 domains, is **targeted** by the Ahi-1 and Mis-2

provirus integrations.

AUTHOR(S): Jiang, Xiaoyan; Hanna, Zaher; Kaouass, Mohammadi; Girard, Luc; Jolicoeur, Paul (1)

CORPORATE SOURCE: (1) Laboratory of Molecular Biology, Clinical Research Institute of Montreal, 110 Pine Ave., West, Montreal, H2W 1R7: jolicop@ircm.qc.ca Canada

SOURCE: Journal of Virology, (September, 2002) Vol. 76, No. 18, pp. 9046-9059. http://intl-jvi.asm.org/. print. ISSN: 0022-538X.

DOCUMENT TYPE: Article

LANGUAGE: English

AB The Ahi-1 locus was initially identified as a common helper provirus integration site in Abelson pre-B-cell lymphomas and shown to be closely linked to the c-myb proto-oncogene. Since no significant alteration of c-myb expression was found in Abelson murine leukemia virus-induced pre-B-lymphomas harboring a provirus inserted within the Ahi-1 locus, this suggested that it harbors another **gene** whose dysregulation is involved in tumor formation. Here we report the identification of a novel **gene** (Ahi-1) **targeted** by these provirus insertional mutations and the cloning of its cDNA. The Ahi-1 proviral insertions were found at the 3' end of the **gene**, in an inverse transcriptional orientation, with most of them located around and downstream of the last exon, whereas another insertion was within intron 22. In addition, another previously identified provirus insertion site, Mis-2, was found to map within the 16th intron of the Ahi-1 **gene**. The Ahi-1 cDNA encodes a 1,047-amino-acid **protein**. The **predicted** Ahi-1 **protein** is a modular **protein** that contains one SH3 motif and seven WD40 **repeats**. The Ahi-1 **gene** is conserved in mammals and encodes two major RNA species of 5 and 4.2 kb and several other shorter splicing variants. The Ahi-1 **gene** is expressed in mouse embryos and in several organs of the mouse and rat, notably, at high levels in the brain and testes. In tumor cells harboring insertional mutations in Ahi-1, truncated Ahi-1/viral fused transcripts were identified, including some splicing variants with deletion of the SH3 domain. Therefore, Ahi-1 is a novel **gene targeted** by provirus insertion and encoding a **protein** that exhibits several **features** of a signaling molecule. Thus, Ahi-1 may play an important role in signal transduction in normal cells and may be involved in tumor development, possibly in cooperation with other oncogenes (such as v-abl and c-myc) or with a tumor suppressor **gene** (Nf1), since Ahi-1 insertion sites were identified in tumors harboring v-abl defective retroviruses or a c-myc transgene or in tumors exhibiting deletion of Nf1.

L13 ANSWER 7 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:604483 PROMT

TITLE: Geron Corporation Announcement: Kyowa Hakko Selects Telomerase Inhibitor Compound for Development.

SOURCE: Business Wire, (13 Aug 2001) pp. 2062.

PUBLISHER: Business Wire

DOCUMENT TYPE: Newsletter

LANGUAGE: English

WORD COUNT: 3568

FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Business Editors and Health/Medical Writers
THIS IS THE FULL TEXT: COPYRIGHT 2001 Business Wire

L13 ANSWER 8 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:798287 PROMT
TITLE: THE MARKET REPORT.(hair care products)(Brief Article)
SOURCE: European Cosmetic Markets, (Oct 2001) Vol. 18, No. 10, pp. 365.
ISSN: 0957-1515.
PUBLISHER: Wilmington Publishing Ltd.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 7922
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Hair styling: To have and to hold
THIS IS THE FULL TEXT: COPYRIGHT 2001 Wilmington Publishing Ltd.

Subscription: 695.00 British pounds per year. Published monthly. Wilmington House, Church Hill, Wilmington, Dartford, Kent DA2 7EF., United Kingdom

L13 ANSWER 9 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:343008 PROMT
TITLE: Rapid progress seen in diagnostic and interventional cardiology.
AUTHOR(S): SIMONSEN, MICHAEL
SOURCE: The BBI Newsletter, (May 2001) Vol. 24, No. 5, pp. 113.
ISSN: 1049-4316.
PUBLISHER: American Health Consultants, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 4245
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB ORLANDO, Florida -- The market for devices used in cardiology is one of the most rapidly growing segments of the medical products market, and continues to offer opportunities for major suppliers as well as new ventures. A wide range of new developments in the cardiovascular device market was presented at the 50th annual scientific sessions of the American College of Cardiology (ACC, Bethesda, Maryland), held here in mid-March.
THIS IS THE FULL TEXT: COPYRIGHT 2001 American Health Consultants, Inc.

Subscription: $725.00 per year. Published monthly.

L13 ANSWER 10 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:750306 PROMT
TITLE: 39th Annual R&D 100 Awards.(Cover Story)(Industry Overview)
SOURCE: R & D, (Sept 2001) Vol. 43, No. 9, pp. 29.
ISSN: 0746-9179.
PUBLISHER: Cahners Business Information
DOCUMENT TYPE: Newsletter
LANGUAGE: English

WORD COUNT: 22673
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB The History Behind the Awards
THIS IS THE FULL TEXT: COPYRIGHT 2001 Cahners Business Information

Subscription: $69.95 per year. Published monthly. 8773 Ridgeline Boulevard, Highlands Ranch, CO 80126-2329.

L13 ANSWER 11 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:378063 PROMT
TITLE: THE KINGS OF MADISON AVENUE.(Brief Article)
SOURCE: Campaign, (27 Apr 2001) pp. S4.
ISSN: 0008-2309.
PUBLISHER: Haymarket Publishing Ltd.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 42255
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB John Wren
THIS IS THE FULL TEXT: COPYRIGHT 2001 Haymarket Publishing Ltd.

Subscription: 2.00 British pounds per year. Published weekly. 22 Lancaster Gate, London W2 3LY., United Kingdom

L13 ANSWER 12 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:184259 PROMT
TITLE: **Genes**, the genome and disease.
AUTHOR(S): Bendall, Kate
SOURCE: New Scientist, (17 Feb 2001) Vol. 169, No. 2278, pp. 1(4).
ISSN: 0262-4079.
PUBLISHER: Reed Elsevier Business Publishing, Ltd.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 3988
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB We have cleared the first hurdle in the race to understand the human genome, but there is much hard work ahead. The potential rewards in medicine, however, are enormous. They include treatments tailor-made to your own **genes** and possible cures for diseases such as cancer and Alzheimer's
THIS IS THE FULL TEXT: COPYRIGHT 2001 Reed Elsevier Business Publishing, Ltd. Subscription: $130.00 per year. Published weekly. Quadrant House, The Quadrant, Sutton, Surrey SM2 5AS., United Kingdom

L13 ANSWER 13 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2002-139705 [18] WPIDS
DOC. NO. CPI: C2002-043008
TITLE: Novel 33358 **polypeptide**, a human ankyrin family member, useful as reagents or **targets** for treating and/or diagnosing myocardial infarction, restenosis, angina, aortic valve stenosis, raynaud's syndrome, psoriasis.
DERWENT CLASS: B04 D16
INVENTOR(S): GLUCKSMANN, M A; KADAMBI, V; KADAMBI, V J

Searcher : Shears 308-4994

PATENT ASSIGNEE(S): (MILL-N) MILLENNIUM PHARM INC; (GLUC-I) GLUCKSMANN M A; (KADA-I) KADAMBI V
COUNTRY COUNT: 96
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2001096375	A2	20011220	(200218)*	EN	115

RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW

AU 2001068570 A 20011224 (200227)
US 2002065223 A1 20020530 (200240)

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2001096375	A2	WO 2001-US19591	20010618
AU 2001068570	A	AU 2001-68570	20010618
US 2002065223	A1 Provisional	US 2000-212222P	20000616
		US 2001-884870	20010618

FILING DETAILS:

PATENT NO	KIND	PATENT NO
AU 2001068570	A Based on	WO 200196375

PRIORITY APPLN. INFO: US 2000-212222P 20000616; US 2001-884870 20010618
AN 2002-139705 [18] WPIDS
AB WO 200196375 A UPAB: 20020319
NOVELTY - An isolated 33358 **polypeptide** (a human ankyrin family member, referred to as cardiac/skeletal muscle-restricted ankyrin-**repeat** containing **protein** (C/SKARP)-1), (I,) comprising a fragment which comprises 10 contiguous amino acids of a fully defined sequence of 323 amino acids (S2), or comprising 60% identity to (S2), is new.

DETAILED DESCRIPTION - Where (I):

(a) comprises a fragment which comprises 10 contiguous amino acids of a fully defined sequence of 323 amino acids (S2);

(b) is a naturally occurring allelic variant of (S2) which is encoded by a **nucleic** acid molecule that hybridizes to a **complement** of a fully defined sequence of 1538 (S1) or 972 (S3) nucleotides under stringent conditions:

(c) is a **polypeptide** which is encoded by a **nucleic** acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleotide sequence of (S1) or (S3); or

(d) is a **polypeptide** comprising an amino acid sequence which is 60% identical to (S2).

INDEPENDENT CLAIMS are also included for the following:

(1) an isolated **nucleic** acid molecule (II);

(2) an isolated **nucleic** acid molecule which hybridizes to a **complement** of (II) under stringent conditions;

(3) an isolated **nucleic** acid molecule comprising a nucleotide sequence which is **complementary** to (II);

(4) an isolated **nucleic** acid molecule comprising (II) and a nucleotide sequence encoding a heterologous **polypeptide**;

(5) a vector (III) comprising (II);

(6) a host cell (IV) transfected with (III);

(7) preparation of (I);

(8) an antibody (V) which selectively binds to (I);

(9) detecting (M1) presence of (II) in a sample;

(10) a kit comprising a compound which selectively binds to (I) or a compound which selectively hybridizes to (II), and instructions for use; and

(11) modulating the activity of (I) involves contacting the **polypeptide** or a cell expressing a **polypeptide** with a compound which binds to the **polypeptide** to modulate the activity of the **polypeptide**.

ACTIVITY - Antiarteriosclerotic; vasotropic; antianginal; cardiant; antiarrhythmic; hypotensive; cytostatic; antitumor; antipsoriatic; antidiabetic; ophthalmological; gynecological; antithyroid; antirheumatic; antiarthritic.

MECHANISM OF ACTION - **Gene** therapy; 33358 **polypeptide** expression or activity modulator.

USE - (I) is useful for identifying a compound which binds to it which involves contacting a **polypeptide** or a cell expressing a **polypeptide** with a test compound and determining whether the **polypeptide** binds to the test compound. The binding of test compound to the **polypeptide** is detected by directly detecting test compound/**polypeptide** binding, or by competition binding assay, or by using an assay for C/SKARP-1 activity. (I) is useful for identifying a compound which modulates its activity which involves contacting (I) with a test compound and determining the effect of the test compound on the activity of the **polypeptide**, to thereby identify a compound which modulates the activity of the **polypeptide**. (IV) is useful for producing (I) by recombinant **techniques**. (V) is useful for detecting the presence of (I) in a sample which involves contacting the sample with a compound (i.e., (V)) which selectively binds to (I) and determining whether (V) binds to the **polypeptide** in the sample (all claimed). The sequences of (I) and (II) are useful as query sequences to perform a search against public databases to identify other family members or related sequences. (I) is useful for treating disorders characterized by insufficient or excessive production of a 33358 substrate or for producing 33358 inhibitors, screening for drugs or compounds which modulate 33358 activity. (I) has the following activities:

(i) mediation of specific macromolecular interactions;

(ii) mediation of interactions between **proteins** and/or between regions of a single **protein**;

(iii) formation of binding sites for distinct **proteins** (e.g., non-C/SKARP **proteins**);

(iv) bridging of cellular components;

(v) regulation of **gene** expression (e.g., cardiac **gene** expression) and, thus, can be used to, for example:

(a) modulate cellular localization (e.g., anchoring C/SKARP

binding **proteins** in a specific cellular localization);

(b) modulate development and/or differentiation (e.g., myogenic development and/or differentiation, heart development and/or differentiation);

(c) modulate cardiac maturation and/or morphogenesis;

(d) as a marker (e.g., an early marker) of cardiac and/or myogenic cell lineage; and

(e) modulate and/or treat C/SKARP-1-associated or related disorders.

(II) is useful for expressing (I) via a recombinant expression vector in a host cell in gene therapy applications, to detect 33358 mRNA, or genetic alteration in 33358 gene and to modulate 33358 activity. Fragments of (II) are useful as probes and primers. Portions or fragments of (II) are useful to:

(i) map their respective genes on a chromosome, e.g., to locate gene regions associated with genetic disease or to associate 33358 with a disease;

(ii) identify an individual from a minute biological sample (tissue typing);

(iii) aid in forensic identification of a biological sample.

(II) can be inserted into vectors and used as gene therapy vectors.

(V) is useful for isolating (I), to detect (I), to evaluate the abundance and pattern of expression of the protein, to diagnostically monitor protein levels in tissue, as a part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. (V) is also useful for isolating 33358 proteins, regulating the bioavailability of 33358 proteins and modulating 33358 activity. (I), (II) and (V) are useful for:

(a) screening assays;

(b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and

(c) methods of treatment (e.g., therapeutic and prophylactic).

33358 molecules ((I), (II), (V)) are useful as a novel diagnostic target and therapeutic agent in assays applicable to treatment and diagnosis of 33358 mediated or related disorders such as cardiac hypertrophy, cardiac disorders and/or cardiovascular disease, (e.g., congestive heart failure, cardiomyopathy), arteriosclerosis, ischemia reperfusion injury, restenosis, tachycardia, bradycardia, angina, hypertension, myocardial infarction, coronary artery disease, ischemic disease, raynaud's syndrome, aneurysm, aortic valve stenosis. A cardiovascular disease or disorder also includes an endothelial cell disorder, and thus the molecules are useful for treating the endothelial cell disorder such as tumorigenesis, tumor metastasis, psoriasis, diabetic retinopathy, endometriosis, Grave's disease, ischemic disease (e.g., atherosclerosis), and chronic inflammatory diseases (e.g., rheumatoid arthritis). The 33358 molecules are useful as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states, as pharmacodyanamic markers and as pharmacogenomic markers.

Dwg.0/3

L13 ANSWER 14 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2001-616539 [71] WPIDS
DOC. NO. CPI: C2001-184682
TITLE: **Gene** expression assay for analyzing **gene** expression, involves providing array

Searcher : Shears 308-4994

of **nucleic** acid mixtures at addressable locations on substrate and exposing array to probe for detecting **nucleic** acid molecules on array.

DERWENT CLASS: B04 D16
INVENTOR(S): MARINCOLA, F M; MILLER, L D; WANG, E
PATENT ASSIGNEE(S): (USSH) US DEPT HEALTH & HUMAN SERVICES; (USSH) US NAT INST OF HEALTH
COUNTRY COUNT: 95
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2001073134	A2	20011004	(200171)*	EN	60

RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW

AU 2001051069 A 20011008 (200208)

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2001073134	A2	WO 2001-US9993	20010328
AU 2001051069	A	AU 2001-51069	20010328

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 2001051069	A	Based on	WO 200173134

PRIORITY APPLN. INFO: US 2000-192700P 20000328
AN 2001-616539 [71] WPIDS
AB WO 200173134 A UPAB: 20011203
NOVELTY - A **gene** expression assay (M1) involves providing an array (A) of **nucleic** acid mixtures at addressable locations on a substrate, where the mixtures comprise polynucleotides (P) in quantities substantially proportional to quantities of polynucleotides in a specimen from which P are obtained, and exposing A to a probe for detecting P on A under conditions sufficient to produce binding of probe to P.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) an assay (M2) to determine relative expression of a **DNA** sequence in a number of biological specimens, involves providing a labeled probe, contacting the labeled probe with an array of mixtures of **nucleic** acid molecules arrayed on a surface of a solid support, under conditions sufficient to produce binding, where each mixture of **nucleic** acid molecules proportionately reflects expression levels of RNA molecules from a specimen from which the **nucleic** acid molecules are obtained, separating unbound labeled probe from the array, and detecting probe binding on the array;

(2) a **gene** profiling array (I) comprises a number of mixtures of **nucleic** acid molecules immobilized on a solid support in an addressable pattern, where each mixture proportionately reflects expression levels of mRNA molecules in a specimen from which the **nucleic** acid molecules are obtained;

(3) a kit (II) for determining relative expression of a **DNA** sequence in a number of biological specimens, comprises (I), and instructions for using (I);

(4) an assay (M3) for analyzing a number of **gene** expression profiles, comprises providing (I), exposing (I) to a first probe that may hybridize to the **nucleic** acid molecules of the array to identify those **nucleic** acid molecules to which the first probe hybridizes, detecting a first hybridization pattern of the first probe, and **repeating** the exposing and detecting steps with a second probe to identify samples to which the second probe hybridizes;

(5) producing (M4) a mixture of mRNA-derived **nucleic** acid molecules, involves isolating an RNA sample from a specimen, obtaining one or more RNA templates from a portion of the RNA sample, hybridizing the templates with a first primer to form a primed template, where the first primer comprises an antisense sequence of an RNA polymerase promoter, synthesizing first strand cDNA from the primed template, hybridizing the first strand cDNA with a second primer to form a switched template, where the second primer has a 5' end and a 3' end and comprises a string of dG residues at the 3' end, synthesizing second strand cDNA from the switched template to generate full-length double stranded cDNA, transcribing antisense (aRNA) from the full-length double stranded cDNA, and reverse transcribing amplified cDNA from the transcribed aRNA; and

(6) an array (III) used in M1.

USE - M2 is useful to determine relative expression of a **DNA** sequence in a number of biological specimens (claimed). M1 or (I) is useful for analyzing **gene** expression, particularly for comparing **gene** expression in a number of cells or tissues simultaneously. M1 is useful to examine progression of **gene** expression changes both in the same and in different tumor types or in diseases other than neoplasia. (I) is useful to evaluate genetic drift, population differences, progressive speciation, and other such evolution-related phenomena, and to track and study genetically-linked diseases such as cancer, polycystic kidney disease, Huntington's disease and hemophilia. (I) is useful to identify and analyze prognostic **markers** or **markers** that **predict** therapy outcome for various diseases or abnormal conditions, such as cancer. (I) is useful for high throughput analysis of differential **gene** expression in a specimen (such as a tumor) or a variety of specimens (such as a variety of tumor), and for automated preparation and analysis.

ADVANTAGE - M4 produces high-fidelity, amplified mixtures of **nucleic** acid molecules using a combination of aRNA amplification and template-switching synthesis. (I) allows high throughput analysis of differential **gene** expression in a specimen (such as a tumor) or a variety of specimen (such as a variety of tumor), and are suitable for automated preparation and analysis. Thousands of different kinds of cells and tissues can be analyzed for **gene** expression simultaneously. In addition, multiple **genes** are simultaneously profiled using probes

labeled with different fluorescent labels. (I) is much more stable than mRNA arrays used for Northern blots, and can be widely applied to laboratory situation without requiring stringent experimental conditions. (I) provides a relatively accurate indication of the level of expression of each **gene** in a cell.
Dwg.0/6

L13 ANSWER 15 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2002-033170 [04] WPIDS
CROSS REFERENCE: 1997-365950 [34]; 2002-673822 [72]
DOC. NO. CPI: C2002-009186
TITLE: Detecting polymorphisms, useful e.g. for analysis of complex polygenic **traits**, by hybridizing **nucleic** acid to **target**-specific and **marker**-specific probes.
DERWENT CLASS: B04 D16
INVENTOR(S): GHANDOUR, G; LIPSHUTZ, R J; SAPOLSKY, R
PATENT ASSIGNEE(S): (AFFY-N) AFFYMETRIX INC
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 6300063	B1	20011009	(200204)*		24

APPLICATION DETAILS:

PATENT NO	KIND		APPLICATION	DATE
US 6300063	B1	CIP of	US 1995-563762	19951129
		Provisional	US 1996-17260P	19960510
			US 1997-853370	19970508

PRIORITY APPLN. INFO: US 1996-17260P 19960510; US 1995-563762 19951129; US 1997-853370 19970508
AN 2002-033170 [04] WPIDS
CR 1997-365950 [34]; 2002-673822 [72]
AB US 6300063 B UPAB: 20021113
NOVELTY - Detecting a polymorphism in a **target nucleic** acid (I) by hybridizing (I) to an array that comprises at least one detection block of probes comprising (I)-specific probes (P1) and two sets of **marker**-specific probes (P2).

DETAILED DESCRIPTION - Detecting a polymorphism in a **target nucleic** acid (I) by hybridizing (I) to an array that comprises at least one detection block of probes comprising (I)-specific probes (P1) and two sets of **marker**-specific probes (P2). P1 are **complementary** to (I) except that they contain many mono-substitutions of positions within n (= 0-5) bases of a base that is **complementary** to a polymorphism, The two sets of P2 are **complementary** to **marker**-specific regions (MSR) upstream and downstream of the **target** sequence, and the two sets differ at single bases, corresponding to known mismatch positions. Polymorphisms are identified by determining hybridization intensities of (I) and MSR.

An INDEPENDENT CLAIM is also included for determining the

genotype of a sample of **nucleic** acid sequences.

USE - The **method** is used to identify, and screen for, polymorphic genetic **markers** (especially single-base polymorphisms), particularly for identifying **genes** responsible for disease-related **traits** (for diagnosis and possibly even treatment), especially for analysis of complex polygenic **traits**.

ADVANTAGE - The **method** provides rapid, automatable and efficient analysis of polymorphisms or biallelic **markers**.

DESCRIPTION OF DRAWING(S) - Diagram illustrating the tiling strategy for polymorphism WI-567 and **predicted** patterns for homozygous and heterozygous forms.
Dwg.3/13

L13 ANSWER 16 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:1091965 PROMT
TITLE: InforMax and Media Cybernetics Announce Distribution Agreement.
SOURCE: PR Newswire, (18 Dec 2000) pp. 8870.
PUBLISHER: PR Newswire Association, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 1201
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB ROCKVILLE, Md. and SILVER SPRING, Md., Dec. 18 /PRNewswire/ -- THIS IS THE FULL TEXT: COPYRIGHT 2000 PR Newswire Association, Inc.

L13 ANSWER 17 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:260755 PROMT
TITLE: Nycomed Amersham Announces Results for the Year Ended 31 December 1999.
SOURCE: PR Newswire, (7 Apr 2000) pp. 8313.
PUBLISHER: PR Newswire Association, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 4523
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB LONDON, Feb. 29 /PRNewswire/ -- THIS IS THE FULL TEXT: COPYRIGHT 2000 PR Newswire Association, Inc.

L13 ANSWER 18 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:107224 PROMT
TITLE: Industry Suppliers.(Product Information)(Brief Article)
SOURCE: Appliance Manufacturer, (Sept 2000) Vol. 48, No. 9, pp. 98.
ISSN: 0003-679X.
PUBLISHER: Business News Publishing Co.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 19268
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Presenting product information from suppliers focused on meeting the component, material and equipment needs of the cross-functional

design teams in the appliance industry.
THIS IS THE FULL TEXT: COPYRIGHT 2000 Business News Publishing Co.

Subscription: $55.00 per year. Published monthly. 755 West Big Beaver Road, P.O. Box 4270 (48099), Troy, MI 48099.

L13 ANSWER 19 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:1110906 PROMT
TITLE: 2000 YEAR IN REVIEW.
AUTHOR(S): Ruggless, Ron
SOURCE: Nation's Restaurant News, (18 Dec 2000) Vol. 34, No. 51, pp. 47.
ISSN: 0028-0518.
PUBLISHER: Lebhar-Friedman, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 18649
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Though some may argue that the new millennium actually will occur this coming Jan. 1, the foodservice industry already ushered in at least one version on New Year's Day 2000.
THIS IS THE FULL TEXT: COPYRIGHT 2000 Lebhar-Friedman, Inc.

Subscription: $89.00 per year. Published weekly. 425 Park Avenue, New York, NY 10022.

L13 ANSWER 20 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:931703 PROMT
TITLE: Immunogenicity The Last Hurdle for Clinically Successful Therapeutic Antibodies.
AUTHOR(S): Adair, Fiona
SOURCE: BioPharm, (Oct 2000) Vol. 13, No. 10, pp. 42.
ISSN: 1040-8304.
PUBLISHER: Advanstar Communications, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 2761
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Monoclonal antibodies have numerous potential applications as therapeutics. As **target**-specific drugs, they should cause fewer side effects than many treatments for cancer and other diseases. However, immunogenicity has been a significant barrier to clinical success. Several routes of molecular engineering are being explored to prevent immune response to MAbs in patients, and the best route has yet to be proven.
THIS IS THE FULL TEXT: COPYRIGHT 2000 Advanstar Communications, Inc.

Subscription: $59.00 per year. Published monthly. 131 West First Street, Duluth, MN 55082.

L13 ANSWER 21 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2001:105724 PROMT
TITLE: ANTEC 2000 WRAPUP.
SOURCE: Plastics Engineering, (July 2000) Vol. 56, No. 7, pp.

27.
PUBLISHER: ISSN: 0091-9578.
DOCUMENT TYPE: Society of Plastics Engineers, Inc.
LANGUAGE: Newsletter
WORD COUNT: English
25237

AB *FULL TEXT IS AVAILABLE IN THE ALL FORMAT*
James Brackeen Becomes SPE president at ANTEC 2000 in Orlando, Florida
THIS IS THE FULL TEXT: COPYRIGHT 2000 Society of Plastics Engineers, Inc.

Subscription: $50.00 per year. Published monthly. 14 Fairfield Drive, P.O. Box 0403, Brookfield, CT 06804-0403.

L13 ANSWER 22 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:877164 PROMT
TITLE: HIPAA RESOURCE GUIDE.
SOURCE: Health Data Management, (Sept 2000) Vol. 8, No. 9, pp. S1.
ISSN: 1079-9869.
PUBLISHER: American Banker-Bond Buyer
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 10861

AB *FULL TEXT IS AVAILABLE IN THE ALL FORMAT*
A Guide to Companies Offering Specialized Technologies and Consulting Services
THIS IS THE FULL TEXT: COPYRIGHT 2000 American Banker-Bond Buyer

Subscription: $98.00 per year. Published monthly.

L13 ANSWER 23 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:990037 PROMT
TITLE: LITERATURE SHOWCASE.
SOURCE: Household & Personal Products Industry, (Oct 2000) Vol. 37, No. 10, pp. A.
ISSN: 0090-8878.
PUBLISHER: Rodman Publications, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 4796

AB *FULL TEXT IS AVAILABLE IN THE ALL FORMAT*
The following pages offer new brochures and product information from companies advertising in this section. To receive a particular brochure or more information on a product, please circle the appropriate number on the Reader Service Card.
THIS IS THE FULL TEXT: COPYRIGHT 2000 Rodman Publications, Inc.

Subscription: $48.00 per year. Published monthly. 17 S. Franklin Turnpike, Box 555, Ramsey, NJ 07446.

L13 ANSWER 24 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2000:454079 PROMT
TITLE: Advertising's Double Helix: A Proposed New Process

Model.(Brief Article)
AUTHOR(S): HUEY, BILL
SOURCE: Journal of Advertising Research, (May 1999) Vol. 39, No. 3, pp. 43.
ISSN: 0021-8499.
PUBLISHER: Advertising Research Foundation, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 4507
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Most process models of how advertising works are linear and posit that receivers of advertising messages move through several hierarchical stages traditionally described as Attention, Interest, Desire, and Action (AIDA); Awareness, Comprehension, Conviction, and Action (DAGMAR); or Knowledge, Liking, Preference, Conviction, and Purchase (Learn-Feel-Do).

Subscription: $100.00 per year. Published bimonthly. 641 Lexington Ave., 11th Floor, New York, NY 10022. FAX 212-319-5265.

L13 ANSWER 25 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 1999:536450 PROMT
TITLE: Executive Roundtable: Drug Discovery.(Industry Overview)
AUTHOR(S): Parle, Elizabeth
SOURCE: Chemical Market Reporter, (16 Aug 1999) Vol. 256, No. 7, pp. FR 19.
ISSN: 1092-0110.
PUBLISHER: Schnell Publishing Company, Inc.
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 4044
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Representatives from genomics and combinatorial chemistry firms provide their views of the current and future direction of drug discovery.

L13 ANSWER 26 OF 33 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2000:27382 BIOSIS
DOCUMENT NUMBER: PREV200000027382
TITLE: IS1626, a new IS900-related Mycobacterium avium insertion sequence.
AUTHOR(S): Puyang, Xiaoling; Lee, Karen; Pawlichuk, Corey; Kunimoto, Dennis Y. (1)
CORPORATE SOURCE: (1) Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, T6G 2H7 Canada
SOURCE: Microbiology (Reading), (Nov., 1999) Vol. 145, No. 11, pp. 3163-3168.
ISSN: 1350-0872.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

AB An insertion sequence designated IS1626 was isolated and characterized from a Mycobacterium avium clinical strain. IS1626 was detected by high-stringency hybridization with the pMB22/S12 probe from IS900 of Mycobacterium paratuberculosis. IS1626 is 1418 bp in size and has a G+C content of 65 mol%. It has neither terminal inverted **repeats** nor flanking direct **repeats**.
Analysis of three IS1626 insertion sites in the M. avium strain the corresponding potential insertion sites in two IS1626-free M. avium strains indicated a consensus sequence of CATGCN(4-5) TCCTN(2) G for IS1626 insertion. In the three clones examined, IS1626 has the same orientation with respect to this **target** site. IS1626 has two major ORFs. ORF1179 encodes a **predicted protein** of 393 amino acids. ORF930, on the **complementary** strand of ORF1179, encodes a **protein** of 310 amino acids. The Shine-Dalgarno sequence for ORF930 is partially located in the flanking region, similar to other IS900-related elements. Analysis of the comparable **features** of insertion sequences and their variable occurrence in related organisms is useful for studying the evolution of these elements and their hosts.

L13 ANSWER 27 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 1998:451720 PROMT
TITLE: Management of H1V infection and other related issues, Part 1
SOURCE: Drug Store News, (24 Aug 1998) pp. CP59.
ISSN: 0191-7587.
LANGUAGE: English
WORD COUNT: 4114
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Introduction
Acquired immunodeficiency syndrome has attracted much attention in the medical field, as well as in the general community. Less than 20 years old, this syndrome has now become a major cause of morbidity and mortality in the United States, particularly among young adults. It is characterized by the development of opportunistic infections and malignancies resulting from a weakened immune system.
AIDS was first described in the United States in 1981, when cases of pneumocystis carinii pneumonia and Kaposi's sarcoma were reported among young homosexual men in New York and California. In 1983, researchers isolated human immunodeficiency virus as the pathogen causing AIDS and serologic tests were developed two years later to detect antibodies to HIV in the blood. This facilitated detection in patients who were clinically asymptomatic and had not yet developed AIDS-related opportunistic infections and donated blood could now be screened. Shortly thereafter, in 1987, the first antiretroviral agent for treating HIV infection was approved by the U.S. Food and Drug [Administration.sup.(1)(2)].
Over the past 15 years, important advances have been made in understanding this syndrome and its clinical management. Advances in technology have made it possible to quantify the number of HIV RNA particles in plasma, which may be useful in **predicting** disease progression and in guiding treatment of HIV infection. The approach to treating HIV infection has also changed, in that combination antiretroviral therapy has been shown superior to monotherapy, which was considered the standard of care several years ago.

Pharmacists are in a unique position to screen for overlapping toxicities and drug interactions in therapeutic regimens characterized by polypharmacy and the well-informed pharmacist can play a key role in the education, treatment, compliance and prevention efforts so crucial for the success of HIV/AIDS therapy.

L13 ANSWER 28 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 97:259759 PROMT
TITLE: Comparability (Bioequivalence) Testing of Monoclonal Antibodies
Monoclonal antibody lots should be subjected to comparability testing to determine its equivalence
AUTHOR(S): Schenerman, Mark A.; Phillips, Kristan
SOURCE: BioPharm, (Apr 1997) pp. 20.
ISSN: 1040-8304.
LANGUAGE: English
WORD COUNT: 1306
FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB Mark A. Schenerman and Kristan Phillips
Comparability refers to the assessment of similarities and differences - whether due to modifications of raw materials, process, or a manufacturing facility - of a product manufactured by two means (1). FDA recognizes that the development of a biological product may require changes in the **method** of manufacture and the location of manufacture before or after biological license application (BLA) approval. Comparability is demonstrated through equivalence of the safety, identity, purity, and potency of product manufactured before and after a modification is made. When developing a comparability testing program, you need to consider the potential impact of the manufacturing change on the product and the ability to assess this impact. Note that from a regulatory perspective the term bioequivalence is generally applied to generic drug products regulated by the Office of Generic Drugs at the Center for Drug Evaluation and Research (CDER).
This concept of comparability represents a new regulatory paradigm for biological products. For many years biologics were considered too complex and our analytical tools too primitive to provide an accurate picture of our final product.

L13 ANSWER 29 OF 33 SCISEARCH COPYRIGHT 2003 ISI (R)
ACCESSION NUMBER: 97:473396 SCISEARCH
THE GENUINE ARTICLE: XE358
TITLE: Langerhans cells in contact allergy and in vitro **predictive** tests.
AUTHOR: Schmitt D (Reprint)
CORPORATE SOURCE: HOP EDOUARD HERRIOT, INSERM, UNITE 346, PAVILLON R, F-69437 LYON 03, FRANCE (Reprint)
COUNTRY OF AUTHOR: FRANCE
SOURCE: REVUE FRANCAISE D ALLERGOLOGIE ET D IMMUNOLOGIE CLINIQUE, (MAY 1997) Vol. 37, No. 3, pp. 243-252.
Publisher: EXPANSION SCI FRANCAISE, 31 BLVD LATOUR MAUBOURG, 75007 PARIS, FRANCE.
ISSN: 0335-7457.
DOCUMENT TYPE: Article; Journal
FILE SEGMENT: CLIN

LANGUAGE: French
REFERENCE COUNT: 27
ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Delayed contact hypersensitivity (DCHS) present two original **features**: the route of administration of the antigen is epicutaneous and the presenting cells are essentially epidermal Langerhans cells. This disease typically occurs after **repeated** contact of the skin with occupational or environmental sensitizing agents. Various cell types, various mediators of inflammation and many cytokines are involved in the regulation of the inflammatory and immunological processes of DCHS. Keratinocytes, the first **target** of contact allergens, play an active role in this process by the production of cytokines, but the pivot cell is the Langerhans cell (LC), a dendritic cell of the epidermis, which is able to transmit the antigenic information of allergens and induced a specific proliferative lymphocyte response to the antigen. DCHS reactions occur in two phases: a first so-called sensitisation phase during which the hapten binds to the LC which then migrates to the proximal lymph nodes, where it transmits the information to T lymphocytes. This lymphocyte activation results in T cell clonal expansion. These cells then return to the circulation. They correspond to memory T cells. The so-called revelation step of DCHS is involved during a subsequent contact of the skin with the same allergen. LC perform the same process. Circulating memory T cells and memory T cells present in the dermis are then restimulated by LC and this activation induces cytokine production, responsible for the inflammatory reaction **characteristic** of DCHS, in which lymphocytes, macrophages and eosinophil polymorphonuclear cells are involved. The early events are still poorly elucidated (penetration of the hapten, binding with epidermal **proteins**, capture by LC), as is processing of the antigen by LC which results in presentation to T cells in the lymph node of an immunogenic substance recognized by a specific receptor. On the other hand, the migration of LC and their primary stimulation function has been more clearly defined. Adherence molecules are involved and, during LC-T cell cooperation, accessory molecules, including those belonging to the B7 family, play an essential role. Homing, preferential guidance of specific T lymphocytes towards the site of penetration of the hapten is starting to be more clearly understood due to the discovery of specific molecules er;pressed by endothelial cells (E-selectins, ICAM, VCAM) and T lymphocytes (CLA). Various proinflammatory or antiinflammatory cytokines (IL1, IL6, IL12) produced in the skin participate In regulation of the inflammatory process. Ultraviolet rays (UV) also have a modulatory role on the production of these cytokines and the cell functions involved. Finally, the causes lending to either an inflammatory reaction or a tolerance situation, i.e. lymphocyte clonal anergy, are still poorly elucidated. Current hypotheses involve cytokines (IL12), lymphocyte subpopulations (TH0, TH1, TH2) and a possible antigen-presenting function of keratinocytes to explain this active situation of non-response or restoration of normality. In vitro reproduction of delayed contact hypersensitivity (DCHS) has two objectives: 1) a better understanding of the: pathophysiological mechanisms, 2) development of laboratory tests allowing **prediction** of the allergenic property of substances applied to human skin.

For this purpose, **predictive methods** have been based, up until recently, on the production of DCHS in animals

!guinea-pigs or mice), However, new regulations concerning the use of laboratory animals require the development of alternative **methods** allowing the in vitro detection of the allergenic properties of a given substance.

We have developed a model of human T lymphocyte sensitization to a hapten by using, as the antigen-presenting cell, cells physiologically involved in this process: normal human epidermal Langerhans cells. Our results show that, in contrast with LC freshly isolated from the epidermis, LC cultured for 2 days in the presence of GM-CSF, then treated with a strong hapten, such as TNP, are able to induce proliferation of autologous human T lymphocytes. The maximum proliferation is obtained on the 5th day of coculture and the response is specific for the allergen. In order to determine whether this in vitro T lymphocyte sensitization **technique** to a hapten can be used as an alternative **predicted method**, other haptens were also used in this model. Our results show that strong allergens induce lymphocyte proliferation in all experiments, while weak allergens induce a proliferative response in a very limited number of experiments. Interestingly, we never obtained a lymphoproliferative response with an irritating agent such as lauryl sodium sulphate, These trials therefore suggest that this in vitro model is able to discriminate strong allergens from weak allergens and irritants. A limitation of this test is that it is unable to detect prohaptens, as, in the presence of paraphenylenediamine (PPDA), substance considered to be Intensely allergenic, a lymphocyte response is obtained with Bandrowski base, one of the degradation products of PPDA. Despite its limitations, this in vitro sensitization model offers interesting prospects in the contest of the development of an alternative **method.** One of the advantages of this model is that it does not appear to overestimate the sensitizing effect of irritants. **Complementary** studies have shown that, in vitro, haptens induce migration of LC in a quantitative model. In vitro **predictive** tests are therefore based on the various capacities of LC: isolation from the epidermis, migration in the presence of allergens, specific primary stimulation of T lymphocytes by an allergen presented by LC. Since 1992, the possibility of producing normal human Langerhans cells entirely in vitro has opened up a new field of investigation designed to use this decisive advantage: the possibility of working with naive LC in increasingly standardized in vitro models.

L13 ANSWER 30 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 96:243153 PROMT
TITLE: Pathogenesis and treatment of rheumatoid arthritis
SOURCE: Drug Topics, (6 May 1996) pp. 130.
ISSN: 0012-6616.
LANGUAGE: English
WORD COUNT: 6403

FULL TEXT IS AVAILABLE IN THE ALL FORMAT

AB ggests the need for earlier treatment. In the absence of severe, fulminant, or life-threatening disease, less toxic drugs, such as auranofin or hydroxychloroquine, should be used first. The next level of drug therapy would include azathioprine, parenteral gold salts, methotrexate, sulfasalazine, or penicillamine. And the last tier of therapy would include the cytotoxic drugs, chlorambucil and cyclophosphamide; cyclosporine; or experimental agents or

procedures.
Traditionally, the pyramidal approach to therapy of RA has been useful. Initially, treatment consists of education, rest, exercise, and patient counseling, combined with basic anti-inflammatory therapy with salicylates or other NSAIDs. Additional therapies, or second-line agents, may be added to this foundation.
Because of the potential destructive impact of RA, therapy has become more aggressive. The exact nature and timing of the added therapies, however, depend on the experience of the physician and the response of the patient.
References are available upon request.
Ahmed H. Hikal is Associate Professor, Pharmaceutics Department, and Research Associate Professor, Research Institute of Pharmaceutical Sciences, at the School of Pharmacy, The University of Mississippi
Ethel M. Hikal is a Relief Pharmacist, South Panola Community Hospital, Batesville, Miss., Tallahatchie General Hospital, Charleston, Miss., and at Wal-Mart Pharmacy, Oxford, Mississippi
NONSTEROIDAL ANTI-INFLAMMATORY DRUGS (NSAIDs)

# NSAID	Half-life (hr)	Daily dosage (mg)
Aspirin	4-15	1000-6000 in 4 doses
Diclofenac sodium	1-2	100-200 in 2 doses
Diflunisal	8-12	500-1000 in 2 doses
Etodolac	7	600-1200 in 1-4 doses
Fenoprofen calcium	2.5	1800-3200 in 3-4 doses
Flurbiprofen	3-4	200-300 in 2-3 doses
Ibuprofen	1-3	1200-3200 in 3-4 doses
Indomethacin	2-5	50-200 in 2-4 doses
Ketoprofen	2-4	150-300 in 3-4 doses
Meclofenamate sodium	2	200-400 in 3-4 doses
Nabumetone	22.5-30	500-2000 in 1-2 doses
Naproxen	12-15	500-1500 in 1-3 doses
Oxaprozin	42-50	1200-1800 single dose
Piroxicam	38-45	30 in a single dose
Sulindac	18	150-300 in 2 doses
Tolmetin sodium	1	1200-2400 in 3 doses

L13 ANSWER 31 OF 33 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 1995-052102 [07] WPIDS

CROSS REFERENCE: 1995-052103 [07]; 1995-052104 [07]
DOC. NO. CPI: C1995-023943
TITLE: Selecting single stranded **nucleic** acid with specific free energy of duplex formation - for identification of optimised flanking or ligand binding sequences, matched primers and relative susceptibility to perturbation.
DERWENT CLASS: B04 D16
INVENTOR(S): BENIGHT, A S; FALDASZ, B D; LANE, M J
PATENT ASSIGNEE(S): (UYNY) UNIV NEW YORK STATE RES FOUND
COUNTRY COUNT: 55
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 9500665	A1	19950105	(199507)*	EN	124
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL OA PT SE					
W: AT AU BB BG BR BY CA CH CN CZ DE DK ES FI GB GE HU JP KE KG KP KR KZ LK LU LV MD MG MN MW NL NO NZ PL PT RO RU SD SE SI SK TJ TT UA UZ VN					
AU 9471739	A	19950117	(199522)		
EP 752007	A1	19970108	(199707)	EN	
R: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE					
US 5593834	A	19970114	(199709)		27
JP 09502341	W	19970311	(199720)		142
CN 1129461	A	19960821	(199751)		
AU 691505	B	19980521	(199832)		
AU 9880911	A	19981022	(199903)		

APPLICATION DETAILS:

PATENT NO	KIND		APPLICATION	DATE
WO 9500665	A1		WO 1994-US6799	19940616
AU 9471739	A		AU 1994-71739	19940616
EP 752007	A1		EP 1994-920751	19940616
			WO 1994-US6799	19940616
US 5593834	A	Cont of	US 1993-78759	19930617
			US 1995-427863	19950426
JP 09502341	W		WO 1994-US6799	19940616
			JP 1995-502945	19940616
CN 1129461	A		CN 1994-192876	19940616
AU 691505	B		AU 1994-71739	19940616
AU 9880911	A	Div ex	AU 1994-71739	19940616
			AU 1998-80911	19980821

FILING DETAILS:

PATENT NO	KIND		PATENT NO
AU 9471739	A	Based on	WO 9500665
EP 752007	A1	Based on	WO 9500665
JP 09502341	W	Based on	WO 9500665
AU 691505	B	Previous Publ.	AU 9471739
		Based on	WO 9500665

PRIORITY APPLN. INFO: US 1994-224840 19940408; US 1993-78759 19930617; US 1995-427863 19950426; US

Searcher : Shears 308-4994

1993-153536 19931117

AN 1995-052102 [07] WPIDS

CR 1995-052103 [07]; 1995-052104 [07]

AB WO 9500665 A UPAB: 19950223

A single stranded **nucleic** acid (I) that when hybridised with its **complement** (II) forms a duplex with a preselected value of a free energy parameter (P) is produced by (1) determining the value of P from the duplex formed by a test **nucleic** acid (A), (2) comparing this with a reference value Pr, (3) if Pr and P have the required relationship then (part of) (A) is selected as (part of) (I). If not, steps (1) and (2) are **repeated** for other test molecules until one having the required relationship is found. Also new is a reaction mixt. comprising first and second primers which when hybridised to their **complements** form duplexes with first and second values of P, provided that P (or the amplification rates) of both primers are about the same.

USE - The **method** is used (1) to identify flanking sequences for a particular site, (2) to provide an optimised ligand binding site, (3) to produce matched primer sets, (4) to **predict** relative susceptibility of a duplex site to perturbation and (5) to construct maps of such susceptibility. Typical applications are prodn. of (I) with defined affinity for ligands that regulate expression of **protein** encoded by (I), to produce primers that amplify separate **targets** at the same rate (for multiplex PCR to allow detection of 2 or more microbial contaminants simultaneously) and to identify **gene** sites susceptible to mutation.

ADVANTAGE - Since P is related to ligand binding, melting temp., **target** affinity, resistance to duplex perturbation etc., (I) can be tailored for specific applications, when the **method** is applied to flanking sequences, properties can be optimised without altering the binding site itself. The matched primers allow false positive amplifications to be detected in PCR (identified by an amplification rate different from that of the added primers).

Dwg.10/11

ABEQ US 5593834 A UPAB: 19970228

A **method** of detecting the presence or absence of a preselected or native **DNA** sequence or sequences, is new, which comprises determining at least one set of **DNA** primers which amplify, or generate amplification products for, at least two different regions from the preselected or native **DNA** sequence at approximately equal rates, by:

iterating through the preselected or native **DNA** sequence to select at least one **DNA** subsequence for use as a potential **DNA** primer or set of **DNA** primers;

calculating free energies DeltaGDo of duplex melting for each potential **DNA** primer, which comprises:

summing free-energy values for hydrogen-bonding and stacking interactions for nucleotide bases constituting each potential **DNA** primer, wherein such free-energy values are predetermined by semi-empirical thermochemical **methods**;

determining calculated composite reaction rates for amplifying the corresponding preselected or native **DNA** sequence using each selected or permuted potential **DNA** primer or set of **DNA** primers by means of the equation:

ln (kII/kI) = (kappa/RT)(DeltaGDoI-DeltaGDoII),

or, if the composite reaction rates are binding limited,

determining calculated binding constants of the corresponding preselected **DNA** sequence by means of the equation:

ln (kII/kI)=(kappa/RT)(DeltaGDoI-DeltaGDoII),

wherein kI and kII are composite reaction rate constants for amplification using potential **DNA** primers I and II, respectively, KI and KII are binding constants for an amplifying reagent binding to the preselected or native **DNA** sequence to which potential **DNA** primers I and II anneal, respectively, R is the universal gas constant, T is absolute temperature, kappa is a proportionality constant or function.

wherein kappa is predetermined in accord with the equations for the amplifying reagent and for set sequence length by relating measured composite rates of amplification or binding with potential **DNA** primers or set of potential **DNA** primers I and II to their respective differences in free energies of melting DeltaGDoI and DeltaGDoII, which comprises:

calculating free energies of melting DeltaGDo for at least two potential **DNA** primers if kappa is a proportionality constant, or at least three if a function, determined in accord with the summing step;

measuring composite rates of amplification or generation of amplification products or binding constants for synthetic or native **DNA** sequences to which the potential **DNA** primers anneal; and

relating the measured composite rates of reaction or binding constants to their respective differences in free energy of melting DeltaGDo in accord with the equations; and

synthesising the potential **DNA** primers by chemical or biochemical **methods**;

measuring composite reaction rates of amplification or generation of amplification products using the potential **DNA** primers by any **method** suited for the purpose;

repeating the iterating, calculating, synthesising and measuring steps to determine at least one additional potential **DNA** primer which amplifies, or generates amplification products for, the preselected or native **DNA** sequence, does not interact with any other primer, and has a composite reaction rate approximately equal to that of the first potential **DNA** primer or set of potential **DNA** primers; and

choosing at least one set of potential **DNA** primers with approximately equal calculated composite reaction rates for amplifying, or generating amplification products for, at least two different regions from the corresponding preselected or native **DNA** sequence using each set of potential **DNA** primers, wherein calculated composite reaction rates fall within a predefined deviation about a mean composite reaction rate; and using at least one set of potential **DNA** primers to detect the presence or absence of a preselected or native **DNA** sequence, which comprises:

combining aliquots of at least one set of **DNA** primers with an analytical unknown sample which may or may not contain a preselected or native **DNA** sequence;

performing the amplifying reaction to generate amplified concentrations or amplification products of the preselected or native **DNA** sequence, if present; and

observing by any suitable **qualitative** or quantitative **method** the presence or absence of, the preselected or native **DNA** sequence, thereby detecting the presence or absence of

the preselected or native **DNA** sequence.
Dwg.0/7

L13 ANSWER 32 OF 33 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1994:299496 BIOSIS
DOCUMENT NUMBER: PREV199497312496
TITLE: Isolation and characterization of IS1181, an insertion sequence from Staphylococcus aureus.
AUTHOR(S): Derbise, Anne; Dyke, Keith G. H.; El Solh, Nevine (1)
CORPORATE SOURCE: (1) Natl. Reference Cent. Staphylococci, Laboratoire des Staphylocoques des Streptocoques, Institut Pasteur, 75724 Paris Cedex 15 France
SOURCE: Plasmid, (1994) Vol. 31, No. 3, pp. 251-264. ISSN: 0147-619X.
DOCUMENT TYPE: Article
LANGUAGE: English
AB The **repeated** nucleotide sequence isolated from a methicillin-resistant Staphylococcus aureus isolate displays the **characteristic features** of an insertion sequence and was named IS1181. It has a size of 1512 bp and consists of a 1359-bp open reading frame that encodes a 439-amino-acid **protein** which is **predicted** to be highly basic and 23-bp terminal inverted **complementary repeated** sequences exhibiting six mismatches. The three copies of IS1181 isolated from distinct parts of the chromosome of S. aureus, BM3121, are flanked at their ends by 8-bp direct **repeats**, suggesting a duplication of the **target** sequence. IS1181 exhibits similarities with IS1165 from Leuconostoc mesenteroides and IS1001 from Bordetella parapertusis. IS1181 was detected in at least two to eight copies in 41 of the 52 S. aureus isolates tested, whereas none of the 26 coagulase-negative staphylococci, 24 streptococci, or 11 enterococci analyzed carried nucleotide sequences hybridizing with IS1181.

L13 ANSWER 33 OF 33 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 93:681937 PROMT
TITLE: Infrared extends into the colour region
SOURCE: European Polymers Paint Colour Journal, (28 Apr 1993) pp. 202. ISSN: 0370-1158.
LANGUAGE: English
WORD COUNT: 1536
FULL TEXT IS AVAILABLE IN THE ALL FORMAT
AB by A G Coventry
Reducing cost of production while maintaining high **quality** standards is the ultimate goal for many manufacturing and indeed service companies. While for product performance, conditions of contract and legislative reasons many complex chemical analyses are carried out to ensure conformity to specification, product appearance is still poorly controlled. Colour disagreement between supplier and customer can lead to rejection of the delivery with considerable disruption to both parties business.
The advent of reasonably priced tri-stimulus colorimeters did much to remove the human dependant nature of colour assessment while the sophistication of the spectrophotometer based colour match **prediction** system enabled accurate interpretation of the customers requirements. The language of colour scales grew as

researchers tried to optimise the data obtained for the widely differing colour measurement requirements of different industries but for many, the essentially laboratory-based instrumentation could not provide the necessary data quickly enough to control colour in real time.
For batch production where processing can be held until colour results are obtained this **method** of control is adequate, but for the true dynamic process, it is less than satisfactory. What is required is instrumentation capable of measuring directly in the process, with minimum disruption to the process. Data should be produced instantaneously with no operator intervention and the instrumentation should be capable of withstanding the harsh process environment. Such systems indeed exist, but for many the capital outlay required to acquire and implement the technology has been unacceptably high. Until now.

(FILE 'HCAPLUS' ENTERED AT 11:24:32 ON 16 APR 2003)
L17 630 SEA FILE=HCAPLUS ABB=ON PLU=ON (MICROARRAY? OR MICRO ARRAY?) AND PREDICT?
L18 41 SEA FILE=HCAPLUS ABB=ON PLU=ON L17 AND COMPLEMENT?
L19 7 SEA FILE=HCAPLUS ABB=ON PLU=ON L18 AND (ALGORITHM OR GSSA)

L20 6 L19 NOT (L1 OR L9)

L20 ANSWER 1 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2003:245118 HCAPLUS
TITLE: Identification of co-regulated genes through Bayesian clustering of **predicted** regulatory binding sites
AUTHOR(S): Qin, Zhaohui S.; McCue, Lee Ann; Thompson, William; Mayerhofer, Linda; Lawrence, Charles E.; Liu, Jun S.
CORPORATE SOURCE: Department of Statistics, Harvard University, Cambridge, MA, 02138, USA
SOURCE: Nature Biotechnology (2003), 21(4), 435-439
CODEN: NABIF9; ISSN: 1087-0156
PUBLISHER: Nature Publishing Group
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The identification of co-regulated genes and their transcription-factor binding sites (TFBS) are key steps toward understanding transcription regulation. In addn. to effective lab. assays, various computational approaches for the detection of TFBS in promoter regions of coexpressed genes have been developed. The availability of complete genome sequences combined with the likelihood that transcription factors and their cognate sites are often conserved during evolution has led to the development of phylogenetic footprinting. The modus operandi of this technique is to search for conserved motifs upstream of orthologous genes from closely related species. The method can identify hundreds of TFBS without prior knowledge of co-regulation or coexpression. Because many of these **predicted** sites are likely to be bound by the same transcription factor, motifs with similar patterns can be put into clusters so as to infer the sets of co-regulated genes,

i.e., the regulons. This strategy utilizes only genome sequence information and is **complementary** to and confirmative of gene expression data generated by **microarray** expts. However, the limited data available to characterize individual binding patterns, the variation in motif alignment, motif width, and base conservation, and the lack of knowledge of the no. and sizes of regulons make this inference problem difficult. We have developed a Gibbs sampling-based Bayesian motif clustering (BMC) **algorithm** to address these challenges. Tests on simulated data sets show that BMC produces many fewer errors than hierarchical and K-means clustering methods. The application of BMC to hundreds of **predicted** .gamma.-proteobacterial motifs correctly identified many exptl. reported regulons, inferred the existence of previously unreported members of these regulons, and suggested novel regulons.

L20 ANSWER 2 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:964607 HCAPLUS
DOCUMENT NUMBER: 138:23176
TITLE: Method for gene expression profiling and kit for determining origin of tumors
INVENTOR(S): Su, Andrew I.; Hampton, Garret M.
PATENT ASSIGNEE(S): IRM LLC, Bermuda
SOURCE: PCT Int. Appl., 70 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002101357	A2	20021219	WO 2002-US18628	20020610

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2001-297277P P 20010610

AB This invention provides methods, kits, and **algorithms** for obtaining mol. signatures of cells based on their gene expression profiles. Devices for carrying out mol. signature anal. of unknown samples are also provided. Thus, mRNA profiling of the 10 most commonly fatal carcinomas coupled with supervised machine learning **algorithms** were used to identify subsets of genes whose expression is uniquely characteristic for each of the 10 carcinomas. These genes were used to accurately **predict** the anat. origin of 75 blinded carcinomas, including metastatic lesions, with up to 95% success rates. This study demonstrates the existence of subsets of genes whose transcription is characteristic of specific carcinomas, despite a wide-ranging appearance of the tumor cells, and illustrates the feasibility of **predicting** the anat.

site of tumor origin in the context of multiple diverse tumor classes.

L20 ANSWER 3 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:745384 HCAPLUS
DOCUMENT NUMBER: 138:117989
TITLE: Primer on medical genomics: Part III: **Microarray** experiments and data analysis
AUTHOR(S): Tefferi, Ayalew; Bolander, Mark E.; Ansell, Stephen M.; Wieben, Eric D.; Spelsberg, Thomas c.
CORPORATE SOURCE: Division of Hematology and Internal Medicine, Mayo Clinic, Rochester, MN, USA
SOURCE: Mayo Clinic Proceedings (2002), 77(9), 927-940
CODEN: MACPAJ; ISSN: 0025-6196
PUBLISHER: Dowden Health Media, Inc.
DOCUMENT TYPE: Journal; General Review
LANGUAGE: English
AB A review. Genomics has been defined as the comprehensive study of whole sets of genes, gene products, and their interactions as opposed to the study of single genes or proteins. **Microarray** technol. is one of many novel tools that are allowing global and high-throughput anal. of genes and gene products. In addn. to an introduction on underlying principles, the current review focuses on the use of both **complementary** DNA and oligodeoxynucleotide **microarrays** in gene expression anal. Genome-wide expts. generate a massive amt. of data points that require systematic methods of anal. to ext. biol. useful information. Accordingly, the current educational communication discusses different methods of data anal., including supervised and unsupervised clustering **algorithms**. Illustrative clin. examples show clin. applications, including (1) identification of candidate genes or pathol. pathways (ie, elucidation of pathogenesis); (2) identification of "new" mol. classes of diseases that may be relevant in disease reclassification, prognostication, and treatment selection (ie, class discovery); and (3) use of expression profiles of known disease classes to **predict** diagnosis and classification of unknown samples (ie, class **prediction**). The current review should serve as an introduction to the subject for clinician investigators, physicians and medical scientists in training, practicing clinicians, and other students of medicine.
REFERENCE COUNT: 54 THERE ARE 54 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 4 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:488347 HCAPLUS
DOCUMENT NUMBER: 137:138386
TITLE: Optimal Approach for Classification of Acute Leukemia Subtypes Based on Gene Expression Data
AUTHOR(S): Cho, Ji-Hoon; Lee, Dongkwon; Park, Jin Hyun; Kim, Kunwoo; Lee, In-Beum
CORPORATE SOURCE: Department of Chemical Engineering, Pohang University of Science and Technology, Pohang, 790-784, S. Korea
SOURCE: Biotechnology Progress (2002), 18(4), 847-854
CODEN: BIPRET; ISSN: 8756-7938

PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The classification of cancer subtypes, which is crit. for successful treatment, has been studied extensively with the use of gene expression profiles from oligonucleotide chips or cDNA **microarrays**. Various pattern recognition methods have been successfully applied to gene expression data. However, these methods are not optimal, rather they are high-performance classifiers that emphasize only classification accuracy. In this paper, we propose an approach for the construction of the optimal linear classifier using gene expression data. Two linear classification methods, linear discriminant anal. (LDA) and discriminant partial least-squares (DPLS), are applied to distinguish acute leukemia subtypes. These methods are shown to give satisfactory accuracy. Moreover, we detd. optimally the no. of genes participating in the classification (a remarkably small no. compared to previous results) on the basis of the statistical significance test. Thus, the proposed method constructs the optimal classifier that is composed of a small size **predictor** and provides high accuracy.

REFERENCE COUNT: 32 THERE ARE 32 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 5 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:420684 HCAPLUS
DOCUMENT NUMBER: 137:16670
TITLE: **Prediction** of compound signature using high density gene expression profiling
AUTHOR(S): Hamadeh, Hisham K.; Bushel, Pierre R.; Jayadev, Supriya; DiSorbo, Olimpia; Bennett, Lee; Li, Leping; Tennant, Raymond; Stoll, Raymond; Barrett, J. Carl; Paules, Richard S.; Blanchard, Kerry; Afshari, Cynthia A.
CORPORATE SOURCE: National Institute of Environmental Health Sciences, Research Triangle Park, NC, 27709, USA
SOURCE: Toxicological Sciences (2002), 67(2), 232-240 CODEN: TOSCF2; ISSN: 1096-6080
PUBLISHER: Oxford University Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB DNA **microarrays**, used to measure the gene expression of thousands of genes simultaneously, hold promise for future application in efficient screening of therapeutic drugs. This will be aided by the development and population of a database with gene expression profiles corresponding to biol. responses to exposures to known compds. whose toxicol. and pathol. endpoints are well characterized. Such databases could then be interrogated, using profiles corresponding to biol. responses to drugs after developmental or environmental exposures. A pos. correlation with an archived profile could lead to some knowledge regarding the potential effects of the tested compd. or exposure. We have previously shown that cDNA **microarrays** can be used to generate chem.-specific gene expression profiles that can be distinguished across and within compd. classes, using clustering, simple correlation, or principal component analyses. In this report, we test the hypothesis that knowledge can be gained regarding the nature of blinded samples, using an initial training

set comprised of gene expression profiles derived from rat liver exposed to clofibrate, Wyeth 14,643, gemfibrozil, or phenobarbital for 24 h or 2 wk of exposure. Highly discriminant genes were derived from our database training set using approaches including linear discriminant anal. (LDA) and genetic **algorithm** /K-nearest neighbors (GA/KNN). Using these genes in the anal. of coded liver RNA samples derived from 24-h, 3-day, or 2-wk exposures to phenytoin, diethylhexylpthalate, or hexobarbital led to successful **prediction** of whether these samples were derived from livers of rats exposed to enzyme inducers or to peroxisome proliferators. This validates our initial hypothesis and lends credibility to the concept that the further development of a gene expression database for chem. effects will greatly enhance the hazard identification processes.

REFERENCE COUNT: 37 THERE ARE 37 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L20 ANSWER 6 OF 6 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:172524 HCAPLUS
DOCUMENT NUMBER: 136:211861
TITLE: Method and system for **predicting** splice variant from DNA chip expression data
INVENTOR(S): Wang, Yixin; Hu, Gang
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 15 pp.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002029113	A1	20020307	US 2001-934083	20010822

PRIORITY APPLN. INFO.: US 2000-226680P P 20000822

AB A system and method to **predict** alternative splicing transcripts using DNA chip expression data as a primary data source are disclosed. The system and method may perform **prediction** of alternative splicing of pre-mRNA that may be used, for example, for regulating eukaryotic gene expression.

(FILE 'MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO, COMPENDEX, PROMT, INSPEC, COMPUSCIENCE' ENTERED AT 11:28:20 ON 16 APR 2003)

L21 30 S L19
L22 29 S L21 NOT (L2 OR L12)
L23 23 DUP REM L22 (6 DUPLICATES REMOVED)

L23 ANSWER 1 OF 23 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.DUPLICATE 1
ACCESSION NUMBER: 2003067763 EMBASE
TITLE: Prioritized selection of oligodeoxyribonucleotide probes for efficient hybridization to RNA transcripts.
AUTHOR: Luebke K.J.; Balog R.P.; Garner H.R.
CORPORATE SOURCE: K.J. Luebke, Center for Biomedical Inventions, Univ. of Texas Southwest. Med. Ctr., Dallas, TX 75390-9185, United States. kevin.luebke@utsouthwestern.edu

SOURCE: Nucleic Acids Research, (15 Jan 2003) 31/2 (750-758).
Refs: 41
ISSN: 0305-1048 CODEN: NARHAD
COUNTRY: United Kingdom
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 022 Human Genetics
029 Clinical Biochemistry
LANGUAGE: English
SUMMARY LANGUAGE: English

AB Only a small fraction of short oligonucleotide probes bind efficiently to **complementary** segments in long RNA transcripts. Technologies such as array-based transcript profiling and antisense control of gene expression would benefit greatly from a method for **predicting** probes that bind well to a given target RNA. To develop an **algorithm** for prioritizing selection of probes, we have analyzed **predicted** thermodynamic parameters for the binding of several large sets of probes to **complementary** RNA transcripts. The binding of five of these sets of probes to their RNA targets has been reported by others. In addition, we have used a method for light-directed synthesis of oligonucleotide arrays that we developed to generate two new arrays of surface-bound probes and measured the binding of these probes to their RNA targets. We considered **predicted** free energies for intramolecular base pairing of the oligonucleotide and its RNA target as well as the **predicted** free energy of intermolecular hybridization of probe and target. We find that a reliable **predictor** of probes that will hybridize significantly with their targeted transcripts is the **predicted** free energy of hybridization minus the **predicted** free energy for intramolecular folding of the probe.

L23 ANSWER 2 OF 23 MEDLINE

ACCESSION NUMBER: 2003149840 IN-PROCESS
DOCUMENT NUMBER: 22552671 PubMed ID: 12627170
TITLE: Identification of co-regulated genes through Bayesian clustering of **predicted** regulatory binding sites.
AUTHOR: Qin Zhaohui S; McCue Lee Ann; Thompson William; Mayerhofer Linda; Lawrence Charles E; Liu Jun S
CORPORATE SOURCE: Department of Statistics, Harvard University, Cambridge, MA 02138.
SOURCE: NATURE BIOTECHNOLOGY, (2003 Apr) 21 (4) 435-9.
Journal code: 9604648. ISSN: 1087-0156.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: IN-PROCESS; NONINDEXED; Priority Journals
ENTRY DATE: Entered STN: 20030401
Last Updated on STN: 20030401

AB The identification of co-regulated genes and their transcription-factor binding sites (TFBS) are key steps toward understanding transcription regulation. In addition to effective laboratory assays, various computational approaches for the detection of TFBS in promoter regions of coexpressed genes have been developed. The availability of complete genome sequences combined with the likelihood that transcription factors and their cognate sites are often conserved during evolution has led to the

development of phylogenetic footprinting. The modus operandi of this technique is to search for conserved motifs upstream of orthologous genes from closely related species. The method can identify hundreds of TFBS without prior knowledge of co-regulation or coexpression. Because many of these **predicted** sites are likely to be bound by the same transcription factor, motifs with similar patterns can be put into clusters so as to infer the sets of co-regulated genes, that is, the regulons. This strategy utilizes only genome sequence information and is **complementary** to and confirmative of gene expression data generated by **microarray** experiments. However, the limited data available to characterize individual binding patterns, the variation in motif alignment, motif width, and base conservation, and the lack of knowledge of the number and sizes of regulons make this inference problem difficult. We have developed a Gibbs sampling-based Bayesian motif clustering (BMC) **algorithm** to address these challenges. Tests on simulated data sets show that BMC produces many fewer errors than hierarchical and K-means clustering methods. The application of BMC to hundreds of **predicted** gamma-proteobacterial motifs correctly identified many experimentally reported regulons, inferred the existence of previously unreported members of these regulons, and suggested novel regulons.

L23 ANSWER 3 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2003:149277 BIOSIS
DOCUMENT NUMBER: PREV200300149277
TITLE: Identification of potential serum markers for tumors with lymph node metastasis based on gene expression **microarray**.
AUTHOR(S): Roh, Mee-Sook (1); Kwon, Hyuk-Chan; Jeong, Jin-Sok (1); Kim, Jae-Seok; Rha, Seo-Hee (1); Lee, Hye-Jeong; Jeong, Jin-Sook (1); Lee, Ki-Nam; Choi, Phil-Jo; Son, Choon-Hee; Yoon, Jin-Han; Shin, Sang-Hun; Kim, Hyo-Jin; Choi, Seok-Ryeol; Lee, Jong-Hoon; Song, Gi-Hoon; Hwang, Tae-Ho
CORPORATE SOURCE: (1) Department of Pathology, Medical College, Dong-A University, Busan, 602-714, South Korea South Korea
SOURCE: Cancer Gene Therapy, (January 2003, 2003) Vol. 10, No. Supplement 1, pp. S40-S41. print.
Meeting Info.: Eleventh International Conference on Gene Therapy of Cancer San Diego, CA, USA December 12-14, 2002
ISSN: 0929-1903.
DOCUMENT TYPE: Conference
LANGUAGE: English

L23 ANSWER 4 OF 23 MEDLINE
ACCESSION NUMBER: 2003072123 MEDLINE
DOCUMENT NUMBER: 22470261 PubMed ID: 12582263
TITLE: Identification and removal of contaminating fluorescence from commercial and in-house printed DNA **microarrays**.
AUTHOR: Martinez M Juanita; Aragon Anthony D; Rodriguez Angelina L; Weber Jose M; Timlin Jerilyn A; Sinclair Michael B; Haaland David M; Werner-Washburne Margaret
CORPORATE SOURCE: Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA.

CONTRACT NUMBER: GM67593 (NIGMS)
HG02262 (NHGRI)
SOURCE: NUCLEIC ACIDS RESEARCH, (2003 Feb 15) 31 (4) e18.
Journal code: 0411011. ISSN: 1362-4962.
PUB. COUNTRY: England: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200303
ENTRY DATE: Entered STN: 20030214
Last Updated on STN: 20030304
Entered Medline: 20030303

AB **Microarray** analysis is a critically important technology for genome-enabled biology, therefore it is essential that the data obtained be reliable. Current software and normalization techniques for **microarray** analysis rely on the assumption that fluorescent background within spots is essentially the same throughout the glass slide and can be measured by fluorescence surrounding the spots. This assumption is not valid if background fluorescence is spot-localized. Inaccurate estimates of background fluorescence under the spot create a source of error, especially for low expressed genes. We have identified spot-localized, contaminating fluorescence in the Cy3 channel on several commercial and in-house printed **microarray** slides. We determined through mock hybridizations (without labeled target) that pre-hybridization scans could not be used to **predict** the contribution of this contaminating fluorescence after hybridization because the change in spot-to-spot fluorescence after hybridization was too variable. Two solutions to this problem were identified. First, allowing 4 h of exposure to air prior to printing on to Corning UltraGAPS slides significantly reduced contaminating fluorescence intensities to approximately the value of the surrounding glass. Alternatively, application of a novel, hyperspectral imaging scanner and multivariate curve resolution **algorithms**, allowed the spectral contributions of Cy3 signal, glass, and contaminating fluorescence to be distinguished and quantified after hybridization.

L23 ANSWER 5 OF 23 PROMT COPYRIGHT 2003 Gale Group

ACCESSION NUMBER: 2002:12319 PROMT
TITLE: Gene Logic and Compugen Partner to Expand Genomic Content Offering.
SOURCE: Business Wire, (8 Jan 2002) pp. 2517.
PUBLISHER: Business Wire
DOCUMENT TYPE: Newsletter
LANGUAGE: English
WORD COUNT: 1020

AB *FULL TEXT IS AVAILABLE IN THE ALL FORMAT*
Business Editors/Health & Medical Writers
THIS IS THE FULL TEXT: COPYRIGHT 2002 Business Wire

L23 ANSWER 6 OF 23 WPIDS (C) 2003 THOMSON DERWENT

ACCESSION NUMBER: 2002-750494 [81] WPIDS
DOC. NO. NON-CPI: N2002-591065
DOC. NO. CPI: C2002-212673

TITLE: Analyzing a nucleic acid sequence for **microarray** applications involving nucleic

acid hybridization, comprises constructing a context functional descriptor.

DERWENT CLASS: B04 D16 T01
INVENTOR(S): BENIGHT, A S; HOPFINGER, A J; PANCOSKA, P; RICCELLI, P V
PATENT ASSIGNEE(S): (BIOI-N) BIOINFORMATICS DNA CODES LLC
COUNTRY COUNT: 97
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002072868	A2	20020919	(200281)*	EN	103

RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002072868	A2	WO 2002-US7439	20020311

PRIORITY APPLN. INFO: US 2001-274598P 20010310
AN 2002-750494 [81] WPIDS
AB WO 200272868 A UPAB: 20021216
NOVELTY - Analyzing a nucleic acid sequence by constructing a context functional descriptor (CFD), is new.
DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:
(1) a method of identifying a CFD component associated with a property of a nucleic acid sequence or a peptide encoded by the nucleic acid;
(2) a method of comparing nucleic acid sequences;
(3) a method of providing or identifying a population of nucleic acid sequences;
(4) a method of analyzing a nucleic acid sequence to determine the Tm involved with introducing a change;
(5) a computer-readable file, having a record which includes an element that identifies a nucleic acid, and an element which described the CFD, or one or more of its components; and
(6) a set of nucleic acids.
USE - The method is useful for **microarray** applications involving nucleic acid hybridization.
ADVANTAGE - The method obviates the need to consider or evaluate, explicitly, order dependent sequence specific interactions (e.g., singlet, nearest-neighbors, and next-nearest-neighbors interactions). Further, it has enhanced **predictive** power over existing analytical methods.
Dwg.0/14

L23 ANSWER 7 OF 23 WPIDS (C) 2003 THOMSON DERWENT
ACCESSION NUMBER: 2002-537459 [57] WPIDS
DOC. NO. NON-CPI: N2002-425577

DOC. NO. CPI: C2002-152411
TITLE: Determining the likelihoood for aneuploidy in a cell type or organism, for example, a deletion or duplication of part of a chromosome or of 10 adjacent genes, comprises using expression profiles.
DERWENT CLASS: B04 D16 S03 T01
INVENTOR(S): HUGHES, T R; MARTON, M J
PATENT ASSIGNEE(S): (ROSE-N) ROSETTA INPHARMATICS INC
COUNTRY COUNT: 2
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
WO 2002044411	A1	20020606	(200257)*	EN	137
W: CA JP					

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
WO 2002044411	A1	WO 2000-US35352	20001222

PRIORITY APPLN. INFO: US 2000-250597P 20001201
AN 2002-537459 [57] WPIDS
AB WO 200244411 A UPAB: 20020906
NOVELTY - Determination of the likelihood of aneuploidy in a cell type or organism (A) by the use of expression profiles.

DETAILED DESCRIPTION - Determination of the likelihood of aneuploidy in a cell type or organism (A) by:

(a) quantifying levels of cellular constituents associated with genes (at least 1000, preferably 50000) mapped to different chromosomes in the genome in a cell of (A);

(b) comparing the quantified levels of cellular constituents associated with the genes mapped to the same chromosome, to the mean quantified levels of the cellular constituents associated with the genes; and

(c) identifying the genes mapped to the same chromosome for which the quantified levels of cellular constituents associated with the genes are the same for each of the genes and are dissimilar to the mean quantified levels of the cellular constituents associated with the genes.

The step (c) indicates that aneuploidy of the same chromosome or its portion is likely to be present in (A).

INDEPENDENT CLAIMS are included for:

(1) detecting the predisposition of (A) to a disease associated with aneuploidy;

(2) determining the likelihood of aneuploidy in (A) by detecting an expression bias shared by genes mapped to a single chromosome or mapped to a chromosomal portion in the cell of (A);

(3) detecting the presence of aneuploidy in (A) by comparing (a1) a profile or a **predicted** profile derived from the first profile to a database comprising several landmark profiles to determine at a landmark profile most similar to the first or **predicted** profile;

(4) diagnosing a disease associated with aneuploidy in (A) using (a1);

(5) determining the likelihood of aneuploidy of a gene in (A) by identifying (b1) a cellular constituent that is a member of a wild-type co-varying cellular constituent set;

(6) correcting a profile of a cell type or organism for aneuploidy of a chromosome or chromosomal segment by determining (c1) the value (m1) of the mean chromosomal offset ratio for several genes mapped to the chromosome or chromosomal segment in (A) and dividing (c2) the mean quantified level of several cellular constituents associated with several genes mapped to the chromosome or chromosomal segment by the value of the mean chromosomal offset ratio;

(7) correcting a profile of a cell type or organism for aneuploidy by determining (d1) the value (m2) of the mean chromosomal offset ratio for several genes associated with several constituents whose mean quantified level is altered by the presence of a gene in (A) having an abnormal copy number, and dividing (d2) the mean quantified level of the several cellular constituents that are altered by the presence of a gene having an abnormal copy number by the value of the mean offset ratio or the step (d2) for several genes;

(8) a computer system for determining the likelihood of aneuploidy in (A), for detecting the predisposition of (A) to a disease associated with aneuploidy and for diagnosing the disease associated with the known aneuploidy in (A) comprising a processor unit and a memory unit;

(9) a computer program product for directing a user computer in a computer-aided determination of the likelihood of aneuploidy in (A) comprising either a computer code for carrying out the step (b) or (c);

(10) a computer system for correcting a profile of (A) for aneuploidy of a chromosome or chromosomal segment comprising processor units and the memory units, where the memory units contain a program, which cause the processor units to execute either the steps (c1) and (c2) or the steps (d1) and (d2), or the step (c2) for several genes, or the step (d2) for several genes;

(11) a kit for detecting the presence of aneuploidy comprising a positionally-addressable array of polynucleotide probes bound to a support and either expression profiles, in electronic or written form, each correlated to a known alteration in copy number of a gene, or a container comprising RNA, or cDNA, of a cell having the known aneuploidy.

USE - The method is used for determining the likelihood of aneuploidy (e.g. a deletion or duplication of part of a chromosome or of 10 adjacent genes) in (A) e.g. a human, for detecting the predisposition of (A) to a disease associated with aneuploidy resulting from:

(i) amplification deletion translocation of a gene, where the disease is cancer (e.g. breast cancer, colon cancer, acute myelogenous leukemia, chronic myelocytic leukemia, acute promyelocytic leukemia, acute nonlymphocytic leukemia, acute monocytic leukemia, acute myelomonocytic leukemia, Burkitt's lymphoma, non-Hodgkin's lymphoma, acute lymphoblastic leukemia, chronic lymphocyticleukemia, myeloproliferative diseases, small cell lung cancer, kidney cancer, uterin cancer, cervical cancer, prostate cancer, bladder cancer, ovarian cancer, liposarcoma, synovial sarcoma, rhabdomyosarcoma, extraskeletal myxoid chondrosarcoma, Ewing's tumor, peripheral neuroepithelioma, testicular and ovarian dysgerminoma, retinoblastoma, Wilms' tumor, neuroblastoma, malignant

melanoma, and mesothelioma);

(ii) chromosome trisomy when the disease is Down syndrome, Edwards syndrome, Patau syndrome, triple X syndrome, Klinefelter syndrome, and 47,XYY syndrome; and

(iii) a deletion of a 20 chromosomes or their portions when the disease is cri du chat syndrome, Wolf-Hirschhorn syndrome, Alagille syndrome, Angelman syndrome, DiGeorge syndrome, Langer-Giedion syndrome, Miller-Dieker syndrome, Prader-Willi syndrome, Rubinstein-Taybi syndrome, Smith Magenis syndrome, Williams syndrome and Turner syndrome (all claimed).

ADVANTAGE - The method is more sensitive and comprehensive than previous methods, thus can detect very small insertions or deletions.

Dwg.0/6

L23 ANSWER 8 OF 23 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.DUPLICATE 2
ACCESSION NUMBER: 2002071349 EMBASE
TITLE: Molecular classification of head and neck squamous cell carcinoma using cDNA **microarrays**.
AUTHOR: Belbin T.J.; Singh B.; Barber I.; Socci N.; Wenig B.; Smith R.; Prystowsky M.B.; Childs G.
CORPORATE SOURCE: G. Childs, Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, United States. childs@aecom.yu.edu
SOURCE: Cancer Research, (2002) 62/4 (1184-1190). Refs: 44 ISSN: 0008-5472 CODEN: CNREA8
COUNTRY: United States
DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 005 General Pathology and Pathological Anatomy
016 Cancer
LANGUAGE: English
SUMMARY LANGUAGE: English

AB Squamous cell carcinomas of the head and neck constitute an anatomically heterogeneous group of neoplasms that share in common a causal association with tobacco and alcohol exposure. The clinical course of these neoplasms is difficult to **predict** based on established prognostic clinicopathological criteria. Given the genetic complexity of head and neck cancers, it is not surprising that correlations with individual genetic abnormalities have also been disappointing. Several authors have suggested that global gene expression patterns can be used to subgroup patients with cancer. Here we report the use of cDNA **microarrays** containing 9216 clones to measure global patterns of gene expression in these neoplasms. We have used a statistical **algorithm** to identify 375 genes, which divide patients with head and neck tumors into two clinically distinct subgroups based on gene expression patterns. Our results demonstrate that gene expression profiling can be used as a **predictor** of outcome.

L23 ANSWER 9 OF 23 MEDLINE DUPLICATE 3
ACCESSION NUMBER: 2002472468 MEDLINE
DOCUMENT NUMBER: 22218945 PubMed ID: 12233926
TITLE: Primer on medical genomics. Part III: **Microarray** experiments and data analysis.
AUTHOR: Tefferi Ayalew; Bolander Mark E; Ansell Stephen M; Wieben Eric D; Spelsberg Thomas C

CORPORATE SOURCE: Division of Hematology and Internal Medicine, Mayo Clinic, Rochester, Minn 55905, USA.
SOURCE: MAYO CLINIC PROCEEDINGS, (2002 Sep) 77 (9) 927-40. Ref: 53
Journal code: 0405543. ISSN: 0025-6196.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals; Space Life Sciences
ENTRY MONTH: 200210
ENTRY DATE: Entered STN: 20020918
Last Updated on STN: 20021029
Entered Medline: 20021021

AB Genomics has been defined as the comprehensive study of whole sets of genes, gene products, and their interactions as opposed to the study of single genes or proteins. **Microarray** technology is one of many novel tools that are allowing global and high-throughput analysis of genes and gene products. In addition to an introduction on underlying principles, the current review focuses on the use of both **complementary** DNA and oligodeoxynucleotide **microarrays** in gene expression analysis. Genome-wide experiments generate a massive amount of data points that require systematic methods of analysis to extract biologically useful information. Accordingly, the current educational communication discusses different methods of data analysis, including supervised and unsupervised clustering **algorithms**. Illustrative clinical examples show clinical applications, including (1) identification of candidate genes or pathological pathways (ie, elucidation of pathogenesis); (2) identification of "new" molecular classes of diseases that may be relevant in disease reclassification, prognostication, and treatment selection (ie, class discovery); and (3) use of expression profiles of known disease classes to **predict** diagnosis and classification of unknown samples (ie, class **prediction**). The current review should serve as an introduction to the subject for clinician investigators, physicians and medical scientists in training, practicing clinicians, and other students of medicine.

L23 ANSWER 10 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
DUPLICATE 4
ACCESSION NUMBER: 2002:453832 BIOSIS
DOCUMENT NUMBER: PREV200200453832
TITLE: An **algorithm** for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation **microarray** experiments.
AUTHOR(S): Liu, X. Shirley; Brutlag, Douglas L.; Liu, Jun S. (1)
CORPORATE SOURCE: (1) Department of Statistics, Harvard University, 1 Oxford Street, Cambridge, MA, 02138: jliu@stat.harvard.edu USA
SOURCE: Nature Biotechnology, (August, 2002) Vol. 20, No. 8, pp. 835-839. http://www.nature.com/nbt/. print.
ISSN: 1087-0156.
DOCUMENT TYPE: Article
LANGUAGE: English

AB Chromatin immunoprecipitation followed by cDNA **microarray** hybridization (ChIP-array) has become a popular procedure for studying genome-wide protein-DNA interactions and transcription regulation. However, it can only map the probable protein-DNA interaction loci within 1-2 kilobases resolution. To pinpoint interaction sites down to the base-pair level, we introduce a computational method, Motif Discovery scan (MDscan), that examines the ChIP-array-selected sequences and searches for DNA sequence motifs representing the protein-DNA interaction sites. MDscan combines the advantages of two widely adopted motif search strategies, word enumeration and position-specific weight matrix updating, and incorporates the ChIP-array ranking information to accelerate searches and enhance their success rates. MDscan correctly identified all the experimentally verified motifs from published ChIP-array experiments in yeast (STE12, GAL4, RAP1, SCB, MCB, MCM1, SFF, and SW15), and **predicted** two motif patterns for the differential binding of Rap1 protein in telomere regions. In our studies, the method was faster and more accurate than several established motif-finding **algorithms**. MDscan can be used to find DNA motifs not only in ChIP-array experiments but also in other experiments in which a subgroup of the sequences can be inferred to contain relatively abundant motif sites. The MDscan web server can be accessed at http://BioProspector.stanford.edu/MDscan/.

L23 ANSWER 11 OF 23 MEDLINE DUPLICATE 5
ACCESSION NUMBER: 2002272784 MEDLINE
DOCUMENT NUMBER: 22008301 PubMed ID: 12011482
TITLE: **Prediction** of compound signature using high density gene expression profiling.
COMMENT: Comment in: Toxicol Sci. 2002 Jun;67(2):155-6
AUTHOR: Hamadeh Hisham K; Bushel Pierre R; Jayadev Supriya; DiSorbo Olimpia; Bennett Lee; Li Leping; Tennant Raymond; Stoll Raymond; Barrett J Carl; Paules Richard S; Blanchard Kerry; Afshari Cynthia A
CORPORATE SOURCE: National Institute of Environmental Health Sciences, P.O. Box 12233, MD2-04, Research Triangle Park, NC 27709, USA.
SOURCE: TOXICOLOGICAL SCIENCES, (2002 Jun) 67 (2) 232-40. Journal code: 9805461. ISSN: 1096-6080.
PUB. COUNTRY: United States
DOCUMENT TYPE: (EVALUATION STUDIES) Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200210
ENTRY DATE: Entered STN: 20020516
Last Updated on STN: 20021011
Entered Medline: 20021010

AB DNA **microarrays**, used to measure the gene expression of thousands of genes simultaneously, hold promise for future application in efficient screening of therapeutic drugs. This will be aided by the development and population of a database with gene expression profiles corresponding to biological responses to exposures to known compounds whose toxicological and pathological endpoints are well characterized. Such databases could then be interrogated, using profiles corresponding to biological responses to drugs after developmental or environmental exposures. A positive

correlation with an archived profile could lead to some knowledge regarding the potential effects of the tested compound or exposure. We have previously shown that cDNA **microarrays** can be used to generate chemical-specific gene expression profiles that can be distinguished across and within compound classes, using clustering, simple correlation, or principal component analyses. In this report, we test the hypothesis that knowledge can be gained regarding the nature of blinded samples, using an initial training set comprised of gene expression profiles derived from rat liver exposed to clofibrate, Wyeth 14,643, gemfibrozil, or phenobarbital for 24 h or 2 weeks of exposure. Highly discriminant genes were derived from our database training set using approaches including linear discriminant analysis (LDA) and genetic **algorithm** /K-nearest neighbors (GA/KNN). Using these genes in the analysis of coded liver RNA samples derived from 24-h, 3-day, or 2-week exposures to phenytoin, diethylhexylpthalate, or hexobarbital led to successful **prediction** of whether these samples were derived from livers of rats exposed to enzyme inducers or to peroxisome proliferators. This validates our initial hypothesis and lends credibility to the concept that the further development of a gene expression database for chemical effects will greatly enhance the hazard identification processes.

L23 ANSWER 12 OF 23 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2002005532 EMBASE
TITLE: Protein therapeutics: Promises and challenges for the 21st century.
AUTHOR: Weng Z.; DeLisi C.
CORPORATE SOURCE: Z. Weng, Biomedical Engineering Dept., Bioinformatics Program, Boston University, Boston, MA 02215, United States. zhiping@bu.edus
SOURCE: Trends in Biotechnology, (1 Jan 2002) 20/1 (29-35). Refs: 104
ISSN: 0167-7799 CODEN: TRBIDM
PUBLISHER IDENT.: S 0167-7799(01)01846-7
COUNTRY: United Kingdom
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 004 Microbiology
022 Human Genetics
LANGUAGE: English
SUMMARY LANGUAGE: English
AB Recent advances in massively parallel experimental and computational technologies are leading to radically new approaches to the early phases of the drug production pipeline. The revolution in DNA **microarray** technologies and the imminent emergence of its analogue for proteins, along with machine learning **algorithms**, promise rapid acceleration in the identification of potential drug targets, and in high-throughput screens for subpopulation-specific toxicity. Similarly, advances in structural genomics in conjunction with in vitro and in silico evolutionary methods will rapidly accelerate the number of lead drug candidates and substantially augment their target specificity. Taken collectively, these advances will usher in an era of **predictive** medicine, which will move medical practice from reactive therapy after disease onset, to proactive prevention.

L23 ANSWER 13 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:199700 BIOSIS

DOCUMENT NUMBER: PREV200200199700
TITLE: From genomes to vaccines: Leishmania as a model.
AUTHOR(S): Almeida, Renata; Norrish, Alan; Levick, Mark; Vetrie, David; Freeman, Tom; Vilo, Jaak; Ivens, Alasdair; Lange, Uta; Stober, Carmel; McCann, Sharon; Blackwell, Jenefer M. (1)
CORPORATE SOURCE: (1) Cambridge Institute for Medical Research, Addenbrooke's Hospital, University of Cambridge School of Clinical Medicine, Hills Road, Wellcome Trust/MRC Building, Cambridge, CB2 2XY: jennie.blackwell@cimr.cam.ac.uk UK
SOURCE: Philosophical Transactions of the Royal Society of London B Biological Sciences, (29 January, 2002) Vol. 357, No. 1417, pp. 5-11. http://www.jstor.org/journals/02643960.html. print. ISSN: 0962-8436.
DOCUMENT TYPE: General Review
LANGUAGE: English

AB The 35 Mb genome of Leishmania should be sequenced by late 2002. It contains approximately 8500 genes that will probably translate into more than 10 000 proteins. In the laboratory we have been piloting strategies to try to harness the power of the genome-proteome for rapid screening of new vaccine candidate. To this end, **microarray** analysis of 1094 unique genes identified using an EST analysis of 2091 cDNA clones from spliced leader libraries prepared from different developmental stages of Leishmania has been employed. The plan was to identify amastigote-expressed genes that could be used in high-throughput DNA-vaccine screens to identify potential new vaccine candidates. Despite the lack of transcriptional regulation that polycistronic transcription in Leishmania dictates, the data provide evidence for a high level of post-transcriptional regulation of RNA abundance during the developmental cycle of promastigotes in culture and in lesion-derived amastigotes of Leishmania major. This has provided 147 candidates from the 1094 unique genes that are specifically upregulated in amastigotes and are being used in vaccine studies. Using DNA vaccination, it was demonstrated that pooling strategies can work to identify protective vaccines, but it was found that some potentially protective antigens are masked by other disease-exacerbatory antigens in the pool. A total of 100 new vaccine candidates are currently being tested separately and in pools to extend this analysis, and to facilitate retrospective bioinformatic analysis to develop **predictive algorithms** for sequences that constitute potentially protective antigens. We are also working with other members of the Leishmania Genome Network to determine whether RNA expression determined by **microarray** analyses parallels expression at the protein level. We believe we are making good progress in developing strategies that will allow rapid translation of the sequence of Leishmania into potential interventions for disease control in humans.

L23 ANSWER 14 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2003:48663 BIOSIS
DOCUMENT NUMBER: PREV200300048663
TITLE: Development and production of an oligonucleotide MuscleChip: Use for validation of ambiguous ESTs.
AUTHOR(S): Borup, Rehannah H. A.; Toppo, Stefano; Chen, Yi-Wen;

CORPORATE SOURCE: Teslovich, Tanya M.; Lanfranchi, Gerolamo; Valle, Giorgio; Hoffman, Eric P. (1) (1) Research Center for Genetic Medicine, Children's National Medical Center, 111 Michigan Avenue N.W, Washington, DC, 20010, USA: rborup@cnmcresearch.org, stefano@cribi.unipd.it, ychen@cnmcresearch.org, tteslovich@cnmcresearch.org, lanfra@cribi.unipd.it, giorgio.valle@unipd.it, ehoffman@cnmcresearch.org USA

SOURCE: BMC Bioinformatics, (October 29 2002) Vol. 3, No. 33 Cited December 10, 2002, pp. No Pagination. http://www.biomedcentral.com/1471-2105. online. ISSN: 1471-2105.

DOCUMENT TYPE: Article

LANGUAGE: English

AB Background: We describe the development, validation, and use of a highly redundant 120,000 oligonucleotide **microarray** (MuscleChip) containing 4,601 probe sets representing 1,150 known genes expressed in muscle and 2,075 EST clusters from a non-normalized subtracted muscle EST sequencing project (28,074 EST sequences). This set included 369 novel EST clusters showing no match to previously characterized proteins in any database. Each probe set was designed to contain 20-32 25 mer oligonucleotides (10-16 paired perfect match and mismatch probe pairs per gene), with each probe evaluated for hybridization kinetics (Tm) and similarity to other sequences. The 120,000 oligonucleotides were synthesized by photolithography and light-activated chemistry on each **microarray**. Results: Hybridization of human muscle cRNAs to this MuscleChip (33 samples) showed a correlation of 0.6 between the number of ESTs sequenced in each cluster and hybridization intensity. Out of 369 novel EST clusters not showing any similarity to previously characterized proteins, we focused on 250 EST clusters that were represented by robust probe sets on the MuscleChip fulfilling all stringent rules. 102 (41%) were found to be consistently "present" by analysis of hybridization to human muscle RNA, of which 40 ESTs (39%) could be genome anchored to potential transcription units in the human genome sequence. 19 ESTs of the 40 ESTs were furthermore computer-**predicted** as exons by one or more than three gene identification **algorithms**. Conclusion: Our analysis found 40 transcriptionally validated, genome-anchored novel EST clusters to be expressed in human muscle. As most of these ESTs were low copy clusters (duplex and triplex) in the original 28,000 EST project, the identification of these as significantly expressed is a robust validation of the transcript units that permits subsequent focus on the novel proteins encoded by these genes.

L23 ANSWER 15 OF 23 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 2001405120 EMBASE

TITLE: Hybridization cross-reactivity within homologous gene families on glass cDNA **microarrays**.

AUTHOR: Evertsz E.M.; Au-Young J.; Ruvolo M.V.; Lim A.C.; Reynolds M.A.

CORPORATE SOURCE: Dr. M.A. Reynolds, Incyte Genomics, 3160 Porter Avenue, Palo Alto, CA 94304, United States. markr@incyte.com

SOURCE: BioTechniques, (2001) 31/5 (1182-1192). Refs: 19 ISSN: 0736-6205 CODEN: BTNQDO

COUNTRY: United States
DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 029 Clinical Biochemistry
LANGUAGE: English
SUMMARY LANGUAGE: English

AB Glass cDNA **microarrays** can be used to profile the expression of thousands of gene targets in a single experiment. However, the potential for hybridization cross-reactivity needs to be considered when interpreting the results. Here, we describe hybridization experiments with a model array representing four distinct functional classes (families): chemokines, cytochrome P-450 isozymes, G proteins, and proteases. The cDNA clones selected for this array exhibited pairwise sequence identities ranging from 55% to 100%, as determined by a homology scoring **algorithm** (LALIGN). Targets for **microarraying** were amplified by PCR and spotted in 4-fold replication for signal averaging. One designated target from each family was further amplified by PCR to incorporate a T7 promoter sequence for the production of synthetic RNA transcripts. These transcripts were used to generate fluorescent hybridization probes by reverse transcription at varying input concentrations. As expected, hybridization signals were highest at the matching target elements. Targets containing less than 80% sequence identity relative to the hybridization probe sequences showed cross-reactivities ranging from 0.6% to 12%. Targets containing greater than 80% identity showed higher cross-reactivities (26%-57%). These cross-reactive signals were analyzed for statistical correlation with the length of sequence overlap, percent sequence identity, and homology score determined by LALIGN. Overall, percent sequence identity was the best **predictor** of hybridization cross-reactivity. These results provide useful guidelines for interpreting glass cDNA **microarray** data.

L23 ANSWER 16 OF 23 MEDLINE
ACCESSION NUMBER: 2001159541 MEDLINE
DOCUMENT NUMBER: 21131740 PubMed ID: 11237012
TITLE: Experimental annotation of the human genome using **microarray** technology.
AUTHOR: Shoemaker D D; Schadt E E; Armour C D; He Y D; Garrett-Engele P; McDonagh P D; Loerch P M; Leonardson A; Lum P Y; Cavet G; Wu L F; Altschuler S J; Edwards S; King J; Tsang J S; Schimmack G; Schelter J M; Koch J; Ziman M; Marton M J; Li B; Cundiff P; Ward T; Castle J; Krolewski M; Meyer M R; Mao M; Burchard J; Kidd M J; Dai H; Phillips J W; Linsley P S; Stoughton R; Scherer S; Boguski M S
CORPORATE SOURCE: Rosetta Inpharmatics, Inc., Kirkland, Washington 98034, USA.
SOURCE: NATURE, (2001 Feb 15) 409 (6822) 922-7. Journal code: 0410462. ISSN: 0028-0836.
PUB. COUNTRY: England: United Kingdom
DOCUMENT TYPE: (EVALUATION STUDIES) Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200103
ENTRY DATE: Entered STN: 20010404 Last Updated on STN: 20010404

Entered Medline: 20010322

AB The most important product of the sequencing of a genome is a complete, accurate catalogue of genes and their products, primarily messenger RNA transcripts and their cognate proteins. Such a catalogue cannot be constructed by computational annotation alone; it requires experimental validation on a genome scale. Using 'exon' and 'tiling' arrays fabricated by ink-jet oligonucleotide synthesis, we devised an experimental approach to validate and refine computational gene **predictions** and define full-length transcripts on the basis of co-regulated expression of their exons. These methods can provide more accurate gene numbers and allow the detection of mRNA splice variants and identification of the tissue- and disease-specific conditions under which genes are expressed. We apply our technique to chromosome 22q under 69 experimental condition pairs, and to the entire human genome under two experimental conditions. We discuss implications for more comprehensive, consistent and reliable genome annotation, more efficient, full-length **complementary** DNA cloning strategies and application to complex diseases.

L23 ANSWER 17 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2002:250099 BIOSIS
DOCUMENT NUMBER: PREV200200250099
TITLE: Gene expression profiles in multiple myeloma.
AUTHOR(S): Magrangeas, Florence (1); Avet-Loiseau, Herve (1); Nasser, Valery; Doriot, Beatrice; Nguyen, Cathy; Houlgatte, Remi; Bataille, Regis (1); Minvielle, Stephane (1)
CORPORATE SOURCE: (1) U463, INSERM, Nantes France
SOURCE: Blood, (November 16, 2001) Vol. 98, No. 11 Part 1, pp. 733a. http://www.bloodjournal.org/. print.
Meeting Info.: 43rd Annual Meeting of the American Society of Hematology, Part 1 Orlando, Florida, USA December 07-11, 2001
ISSN: 0006-4971.
DOCUMENT TYPE: Conference
LANGUAGE: English

AB Despite high-dose chemotherapy, bisphosphonates and thalidomide, multiple myeloma (MM) remains an incurable malignancy. Possible ways to improve survival, and ideally to achieve cure, are to identify patients at an earlier stage of the disease, and/or to identify new therapeutic targets enabling the development of novel therapeutic strategies. Gene-expression profiling using **complementary** DNA (cDNA) **microarrays** allowed analysis of thousand markers and has been successfully used to identify unrecognized subgroups of a variety of cancers including breast tumors, melanoma and diffuse large B-cell lymphoma. In this study we examined gene expression pattern in 53 patients with MM, 4 with plasma cell leukemia (PCL), 9 with smoldering myeloma (SMM), 6 with monoclonal gammopathy of undetermined significance (MGUS) and in 3 normal plasma-cell specimens. To help for the interpretation of the variation in the tumor samples we also established gene profiling of 2 human myeloma cell lines (HMCL) cultured in five different conditions. Overall, we analyzed 84 samples of either normal or malignant highly purified CD138+ plasma cells using nylon DNA **microarray** consisting of 7200 human genes. Hierarchical clustering **algorithm** has been applied to identify common patterns of gene expression. Genes were grouped on the basis of

similarity in their expression pattern with overall experimental samples. Similarly, samples were grouped on the basis of similarity of their expression pattern of all the genes. The **algorithm** used segregated, with one exception premalignant and normal cells from malignant bone marrow plasma cells, indicating that clusters of coexpressed genes can define a "premalignant" signature. Multidimensional scaling analysis using the 42 genes that define the premalignant signature clearly separated samples in three groups: one group containing premalignant and normal plasma cells, and two subclasses of MM. The gene expression pattern in SMM, MM and PCL is complicated. We started by investigating whether gene expression signature can **predict** response to treatment. We focused on one pattern and reclustered the MM cases using the selected genes. We obtained two branches that accurately distinguished patients according to the achievement or not of complete response after one course of high-dose melphalan. Our results show that marked heterogeneity observed among MM patients (regarding biological and clinical presentation, response to treatment, or survival) can be understood as a set of different biological features that can be separated by global gene expression profiling.

L23 ANSWER 18 OF 23 MEDLINE
ACCESSION NUMBER: 2001325085 MEDLINE
DOCUMENT NUMBER: 21224149 PubMed ID: 11327501
TITLE: A computational neural approach to support the discovery of gene function and classes of cancer.
AUTHOR: Azuaje F
CORPORATE SOURCE: Artificial Intelligence Group and Centre for Health Informatics, Trinity College, Dublin, Ireland.
SOURCE: IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, (2001 Mar) 48 (3) 332-9.
Journal code: 0012737. ISSN: 0018-9294.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200106
ENTRY DATE: Entered STN: 20010611
Last Updated on STN: 20010611
Entered Medline: 20010607

AB Advances in molecular classification of tumours may play a central role in cancer treatment. Here, a novel approach to genome expression pattern interpretation is described and applied to the recognition of B-cell malignancies as a test set. Using cDNA **microarrays** data generated by a previous study, a neural network model known as simplified fuzzy ARTMAP is able to identify normal and diffuse large B-cell lymphoma (DLBCL) patients. Furthermore, it discovers the distinction between patients with molecularly distinct forms of DLBCL without previous knowledge of those subtypes.

L23 ANSWER 19 OF 23 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 2001244966 EMBASE
TITLE: Cluster analysis and promoter modelling as bioinformatics tools for the identification of target genes from expression array data.
AUTHOR: Werner T.
CORPORATE SOURCE: T. Werner, Genomatix Software GmbH, Karlstrasse 55,

SOURCE: D-80333 Munich, Germany. werner@gsf.de
Pharmacogenomics, (2001) 2/1 (25-36).
Refs: 39
ISSN: 1462-2416 CODEN: PARMFL

COUNTRY: United Kingdom
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 022 Human Genetics
029 Clinical Biochemistry

LANGUAGE: English
SUMMARY LANGUAGE: English

AB Expression arrays yield enormous amounts of data linking genes, via their cDNA sequences, to gene expression patterns. This now allows the characterisation of gene expression in normal and diseased tissues, as well as the response of tissues to the application of therapeutic reagents. Expression array data can be analysed with respect to the underlying protein sequences, which facilitates the precise determination of when and where certain groups of genes are expressed. More recent developments of clustering **algorithms** take additional parameters of the experimental set-up into account, focusing more directly on co-regulated set of genes. However, the information concerning transcriptional regulatory networks responsible for the observed expression patterns is not contained within the cDNA sequences used to generate the arrays. Regulation of expression is determined to a large extent by the promoter sequences of the individual genes (and/or enhancers). The complete sequence of the human genome now provides the molecular basis for the identification of many regulatory regions. Promoter sequences for specific cDNAs can be obtained reliably from genomic sequences by exon mapping. In the many cases in which cDNAs are 5'-incomplete, high quality promoter **prediction** tools can be used to locate promoters directly in the genomic sequence. Once sufficient numbers of promoter sequences have been obtained, a comparative promoter analysis of the co-regulated genes and groups of genes can be applied in order to generate models describing the higher order levels of transcription factor binding site organisation within these promoter regions. Such modules represent the molecular mechanisms through which regulatory networks influence gene expression, and candidates can be determined solely by bioinformatics. This approach also provides a powerful alternative for elucidating the functional features of genes with no detectable sequence similarity, by linking them to other genes on the basis of their common promoter structures.

L23 ANSWER 20 OF 23 EMBASE

ACCESSION NUMBER: 2001000479 EMBASE
TITLE: Gene expression profiling of primary breast carcinomas using arrays of candidate genes.
AUTHOR: Bertucci F.; Houlgatte R.; Benziane A.; Granjeaud S.; Adelaide J.; Tagett R.; Loroid B.; Jacquemier J.; Viens P.; Jordan B.; Birnbaum D.; Nguyen C.
CORPORATE SOURCE: D. Birnbaum, Laboratoire d'Oncologie Moleculaire, U119 INSERM, IFR57, 27 Boulevard Lei Roure, 13009 Marseille, France. birnbaum@marseille.inserm.fr
SOURCE: Human Molecular Genetics, (12 Dec 2000) 9/20 (2981-2991).
Refs: 46
ISSN: 0964-6906 CODEN: HMGEE5
COUNTRY: United Kingdom

DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 016 Cancer
022 Human Genetics
LANGUAGE: English
SUMMARY LANGUAGE: English
AB Breast cancer is characterized by an important histoclinical heterogeneity that currently hampers the selection of the most appropriate treatment for each case. This problem could be solved by the identification of new parameters that better **predict** the natural history of the disease and its sensitivity to treatment. A large-scale molecular characterization of breast cancer could help in this context. Using cDNA arrays, we studied the quantitative mRNA expression levels of 176 candidate genes in 34 primary breast carcinomas along three directions: comparison of tumor samples, correlations of molecular data with conventional histoclinical prognostic features and gene correlations. The study evidenced extensive heterogeneity of breast tumors at the transcriptional level. A hierarchical clustering **algorithm** identified two molecularly distinct subgroups of tumors characterized by a different clinical outcome after chemotherapy. This outcome could not have been **predicted** by the commonly used histoclinical parameters. No correlation was found with the age of patients, tumor size, histological type and grade. However, expression of genes was differential in tumors with lymph node metastasis and according to the estrogen receptor status; ERBB2 expression was strongly correlated with the lymph node status ($P < 0.0001$) and that of GATA3 with the presence of estrogen receptors ($P < 0.001$). Thus, our results identified new ways to group tumors according to outcome and new potential targets of carcinogenesis. They show that the systematic use of cDNA array testing holds great promise to improve the classification of breast cancer in terms of prognosis and chemosensitivity and to provide new potential therapeutic targets.

L23 ANSWER 21 OF 23 MEDLINE
ACCESSION NUMBER: 2000210697 MEDLINE
DOCUMENT NUMBER: 20210697 PubMed ID: 10749139
TITLE: Identification of differentially expressed genes in human prostate cancer using subtraction and **microarray**.
AUTHOR: Xu J; Stolk J A; Zhang X; Silva S J; Houghton R L; Matsumura M; Vedvick T S; Leslie K B; Badaro R; Reed S G
CORPORATE SOURCE: Corixa Corporation, Seattle, Washington 98104, USA.
CONTRACT NUMBER: CA80518 (NCI)
SOURCE: CANCER RESEARCH, (2000 Mar 15) 60 (6) 1677-82.
Journal code: 2984705R. ISSN: 0008-5472.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200005
ENTRY DATE: Entered STN: 20000512
Last Updated on STN: 20000512
Entered Medline: 20000504
AB We have identified human prostate cancer- and tissue-specific genes using cDNA library subtraction in conjunction with high throughput **microarray** screening. Subtracted cDNA libraries of prostate tumors and normal prostate tissue were generated. Characterization

of subtracted libraries showed enrichment of both cancer- and tissue-specific genes. Highly redundant clones were eliminated by colony hybridization. The remaining clones were selected for **microarray** to determine gene expression levels in a variety of tumor and normal tissues. Clones showing overexpression in prostate tumors and/or normal prostate tissues were selected and sequenced. Here we report the identification of two genes, P503S and P504S, from subtracted libraries and a third gene, P510S, by subtraction followed by **microarray** screening. Their expression profiles were further confirmed by Northern blot, real-time PCR (TaqMan), and immunohistochemistry to be overexpressed in prostate tissues and/or prostate tumors. Full-length cDNA sequences were cloned, and their subcellular locations were **predicted** by a bioinformatic **algorithm**, PSORT, to be plasma membrane proteins. The genes identified through these approaches are potential candidates for cancer diagnosis and therapy.

L23 ANSWER 22 OF 23 SCISEARCH COPYRIGHT 2003 ISI (R)
ACCESSION NUMBER: 2001:111676 SCISEARCH
THE GENUINE ARTICLE: 394ZG
TITLE: Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells
AUTHOR: Burczynski M E; McMillian M; Ciervo J; Li L; Parker J B; Dunn R T; Hicken S; Farr S; Johnson M D (Reprint)
CORPORATE SOURCE: Robert Wood Johnson Pharmaceut Res Inst, Drug Safety Evaluat, POB 300, Route 202, Raritan, NJ 08869 USA (Reprint); Robert Wood Johnson Pharmaceut Res Inst, Drug Safety Evaluat, Raritan, NJ 08869 USA; Phase 1 Mol Toxicol Inc, Santa Fe, NM 87505 USA
COUNTRY OF AUTHOR: USA
SOURCE: TOXICOLOGICAL SCIENCES, (DEC 2000) Vol. 58, No. 2, pp. 399-415.
Publisher: OXFORD UNIV PRESS, GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND.
ISSN: 1096-6080.
DOCUMENT TYPE: Article; Journal
LANGUAGE: English
REFERENCE COUNT: 38

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB The rapid discovery of sequence information from the Human Genome Project has exponentially increased the amount of data that can be retrieved from biomedical experiments. Gene expression profiling, through the use of **microarray** technology, is rapidly contributing to an improved understanding of global, coordinated cellular events in a variety of paradigms. In the field of toxicology, the potential application of toxicogenomics to indicate the toxicity of unknown compounds has been suggested but remains largely unsubstantiated to date. A major supposition of toxicogenomics is that global changes in the expression of individual mRNAs (i.e., the transcriptional responses of cells to toxicants) will be sufficiently distinct, robust, and reproducible to allow discrimination of toxicants from different classes. Definitive demonstration is still lacking for such specific "genetic fingerprints," as opposed to nonspecific general stress responses that may be indistinguishable between compounds and therefore not suitable as probes of toxic mechanisms. The present studies

demonstrate a general application of toxicogenomics that distinguishes two mechanistically unrelated classes of toxicants (cytotoxic anti-inflammatory drugs and DNA-damaging agents) based solely upon a cluster-type analysis of genes differentially induced or repressed in cultured cells during exposure to these compounds. Initial comparisons of the expression patterns for 100 toxic compounds, using all similar to 250 genes on a DNA **microarray**(similar to 2.5 million data points), failed to discriminate between toxicant classes. A major obstacle encountered in these studies was the lack of reproducible gene responses, presumably due to biological variability and technological limitations. Thus multiple replicate observations for the prototypical DNA damaging agent, cisplatin, and the non-steroidal anti-inflammatory drugs (NSAIDs) diflunisal and flufenamic acid were made, and a subset of genes yielding reproducible inductions/repressions was selected for comparison. Many of the "fingerprint genes" identified in these studies were consistent with previous observations reported in the literature (e.g., the well-characterized induction by cisplatin of p53-regulated transcripts such as p21(waf1/cip1) and PCNA [proliferating cell nuclear antigen]). These gene subsets not only discriminated among the three compounds in the learning set but also showed **predictive** value for the rest of the database (similar to 100 compounds of various toxic mechanisms). Further refinement of the clustering strategy, using a computer-based optimization **algorithm**, yielded even better results and demonstrated that genes that ultimately best discriminated between DNA damage and NSAIDs were involved in such diverse processes as DNA repair, xenobiotic metabolism, transcriptional activation, structural maintenance, cell cycle control, signal transduction, and apoptosis. The determination of genes whose responses appropriately group and dissociate anti-inflammatory versus DNA-damaging agents provides an initial paradigm upon which to build for future, higher throughput-based identification of toxic compounds using gene expression patterns alone.

L23 ANSWER 23 OF 23 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:299311 BIOSIS
DOCUMENT NUMBER: PREV200100299311
TITLE: Systematic analysis and display of oligonucleotide **microarray** based global gene expression data with SPSS 8.0 software.
AUTHOR(S): Zhan, F. (1); Zheng, M. (1); Tricot, G. (1); Barlogie, B. (1); Shaughnessy, J. (1)
CORPORATE SOURCE: (1) Donna and Donald Lambert Laboratory of Myeloma Genetics and Myeloma and Transplantation Research Center, University of Arkansas Medical School, Little Rock, AR USA
SOURCE: Blood, (November 16, 2000) Vol. 96, No. 11 Part 2, pp. 170b. print.
Meeting Info.: 42nd Annual Meeting of the American Society of Hematology San Francisco, California, USA December 01-05, 2000 American Society of Hematology . ISSN: 0006-4971.
DOCUMENT TYPE: Conference
LANGUAGE: English
SUMMARY LANGUAGE: English
AB Oligonucleotide array technologies have made it possible to monitor

simultaneously the expression pattern of thousands of genes. It is of interest to develop techniques for extracting useful information from the resulting data sets. Here we report the application of SPSS 8.0 software for analyzing a data set consisting of expression patterns of 31 multiple myeloma and 7 normal tonsilar derived plasma cell samples. RNA was isolated from highly purified cells (>90% pure) converted to cRNA and hybridized to the HuGeneFL **microarray** evaluating, in a quantitative fashion, 6,800 genes (Affymetrix, Santa Clara CA). Normalized data was subjected to analysis with the t-test or Student t-test to determine whether if the groups were different. Two-way clustering **algorithm**, applied to both the genes and abnormal verses normal cell type, revealed broad coherent patterns that suggests a high degree of organization underlying the gene expression. The clustered gene expression data groups were used to identify genes of known or similar function, tissue specificity and similar subtypes of myeloma. Previous studies have suggested that the transcription factor NF-kB regulates the expression of MB-1. A correlation analysis to study the relationship of NF-kB and MB-1 expression in plasma cells from the 31 MM patients and 7 tonsils resulted in a Pearson correlation coefficient value of of 0.27 and 0.25, respectively (>0.7 being significant). Theses data indicated that no correlation exists between expression of NF-kB and MB-1 in myeloma or normal tonsilar plasma cells. It is possible that correlations will exists between novel genes and genes known to be regulated by NF-kB, a gene important in B-cell development and possibly plasma cell neoplasia. Regression analysis making use of the correlation among gene expression levels was used to establish a **prediction** equation in which independent variables were each assigned a weight based on their relationship to the dependent. An automatically derived class **predictor** was able to determine the class of new MM cases. These results demonstrate that the SPSS-8.0 software can be used to organize genes into biologically relevant clusters based on expression profiles and suggest a general strategy for the dissection of the genetic changes in myeloma and possible class discovery.

FILE 'HOME' ENTERED AT 11:29:57 ON 16 APR 2003